

Mining (and Modeling) Networks to

study

Rare
Common

Diseases



Anaïs Baudot

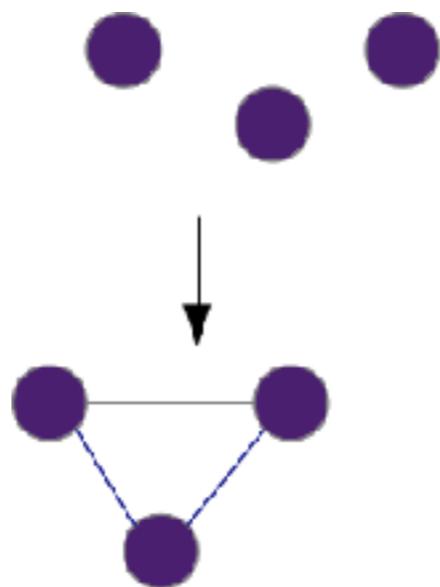
anais.baudot@univ-amu.fr

I2M - UMR 7373

Journées Bioss
3/14, 2017



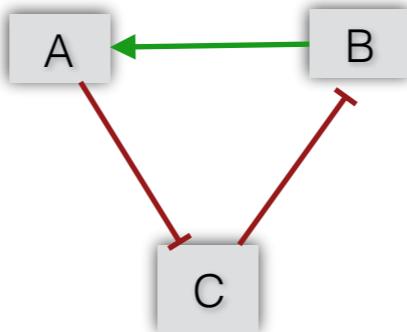
Genotype



Networks !!



Phenotype

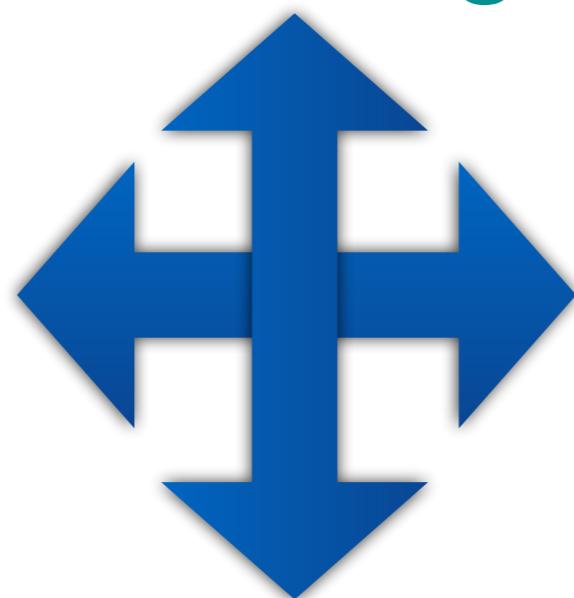


MABIOS

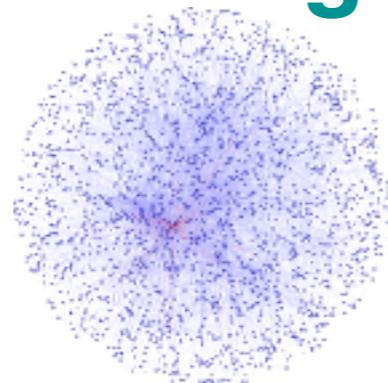
Dynamical Network Modeling

**Algorithmic
developments**

Applications

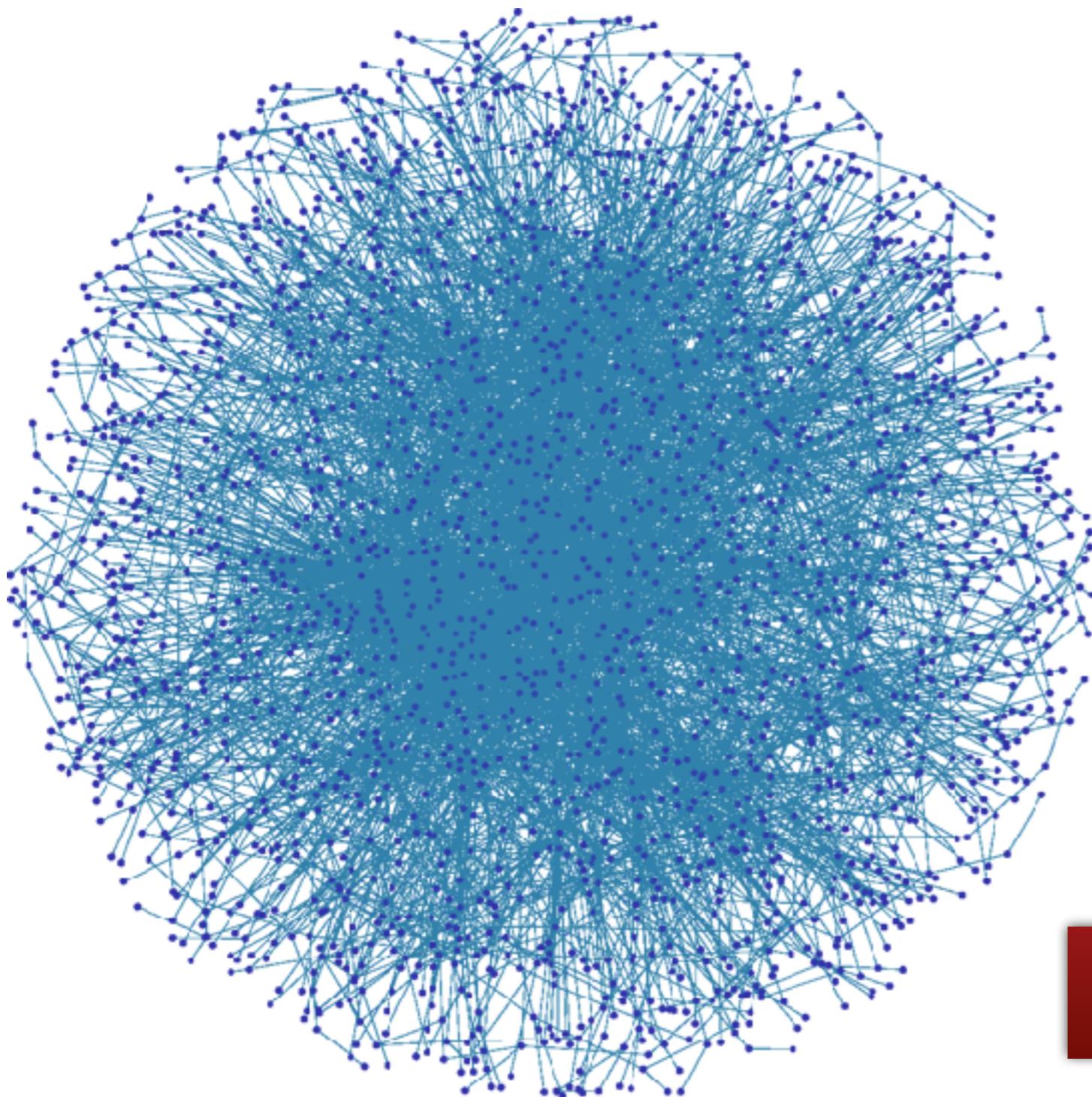


Large-scale Network Mining



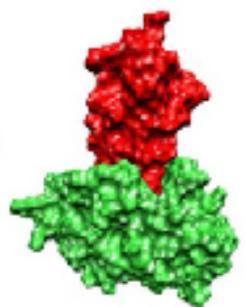


Interactome: The set of possible protein-protein interactions

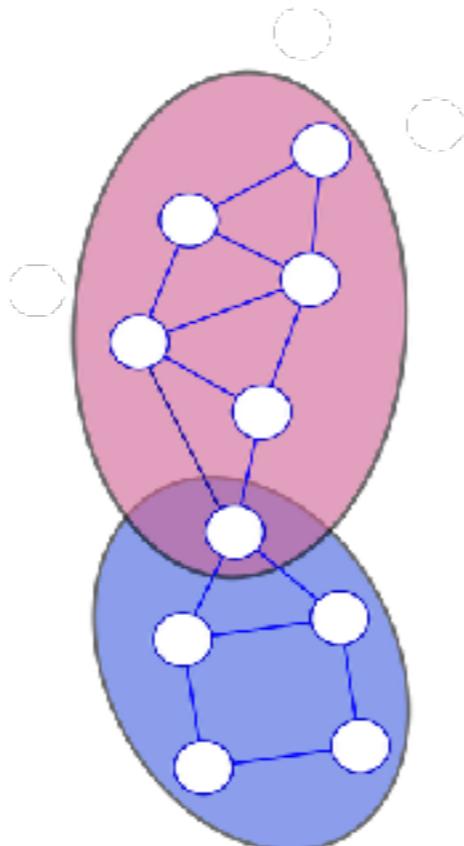


~ 60 000 binary undirected interactions
~ 12 000 human proteins

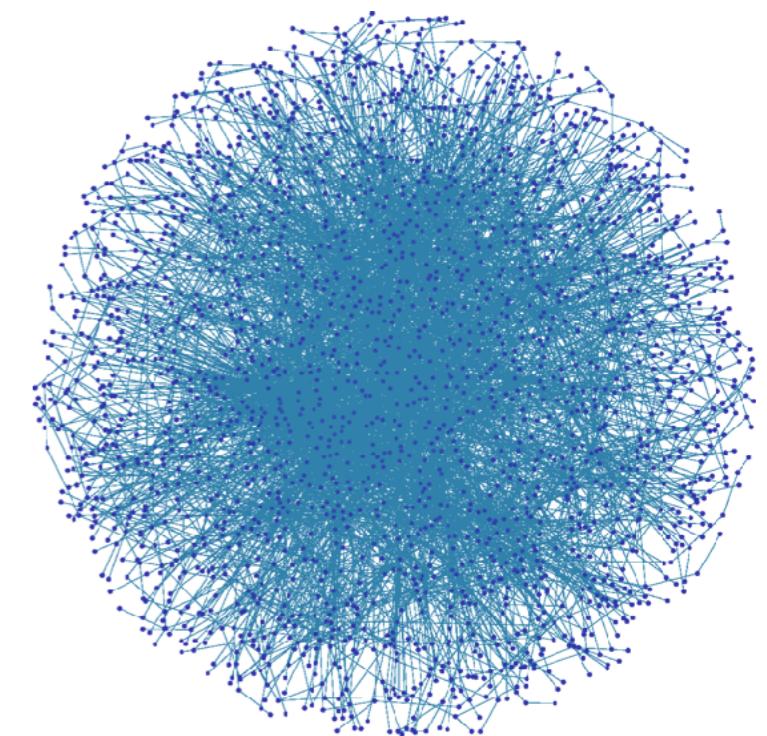
How to extract relevant biological information ?



Binary interaction



Functional modules



Interaction Networks

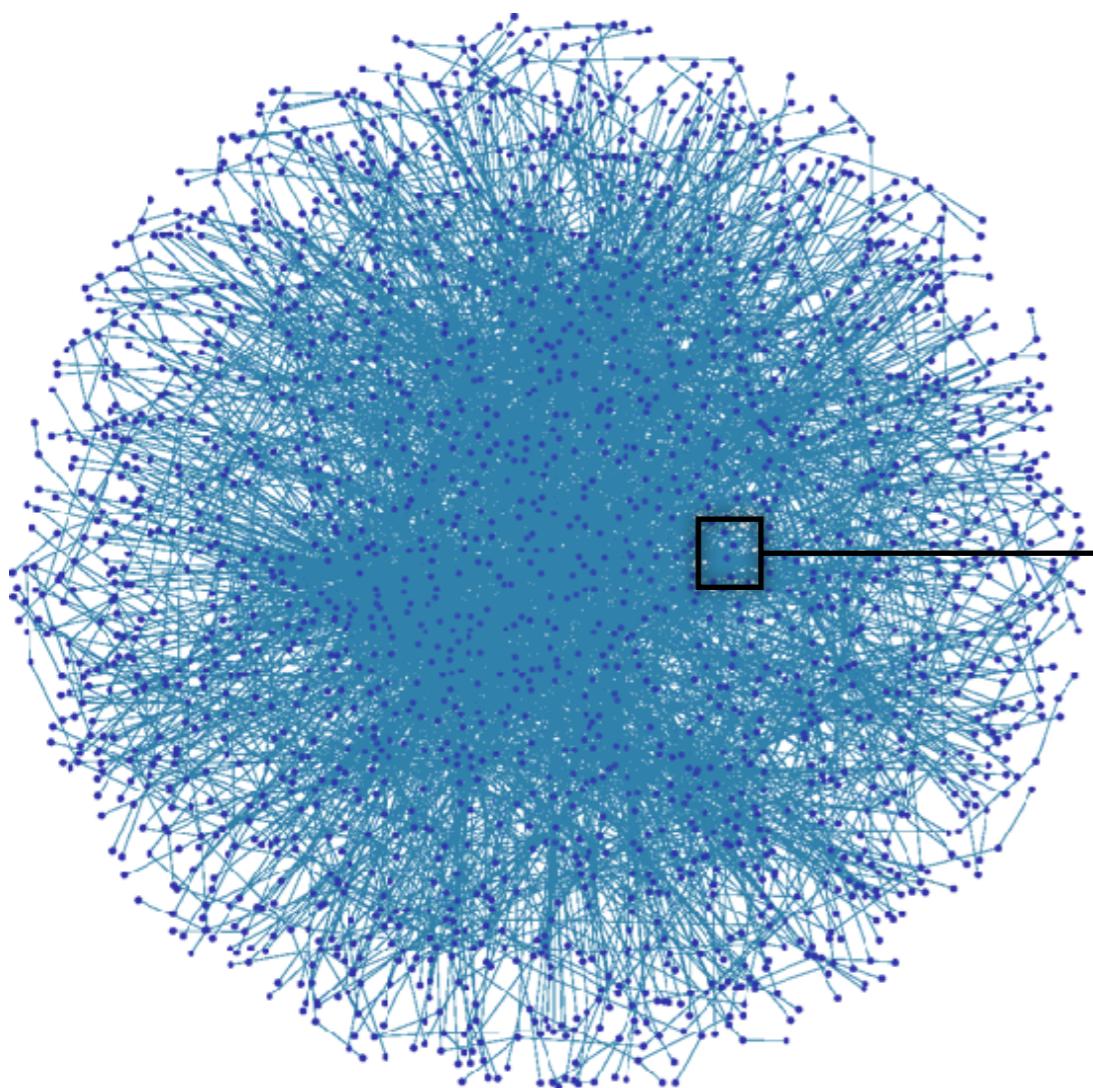
Precision

Comprehensiveness

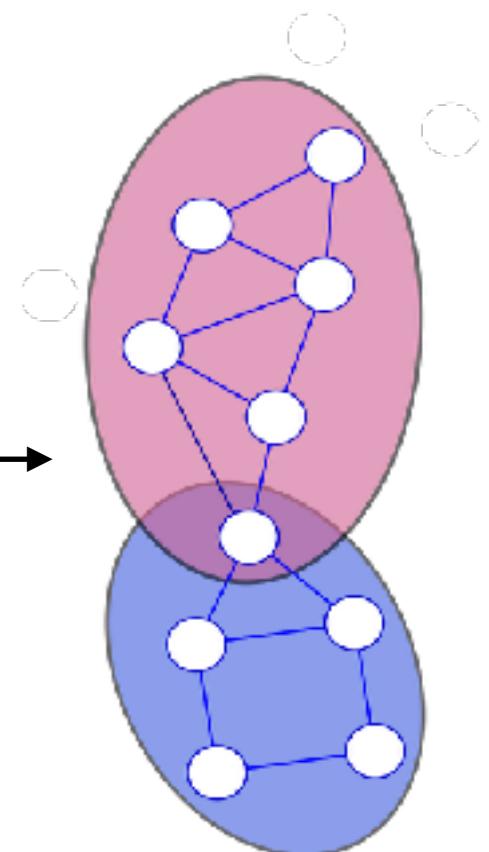
Inspired from P. Aloy,
ECCB 2014

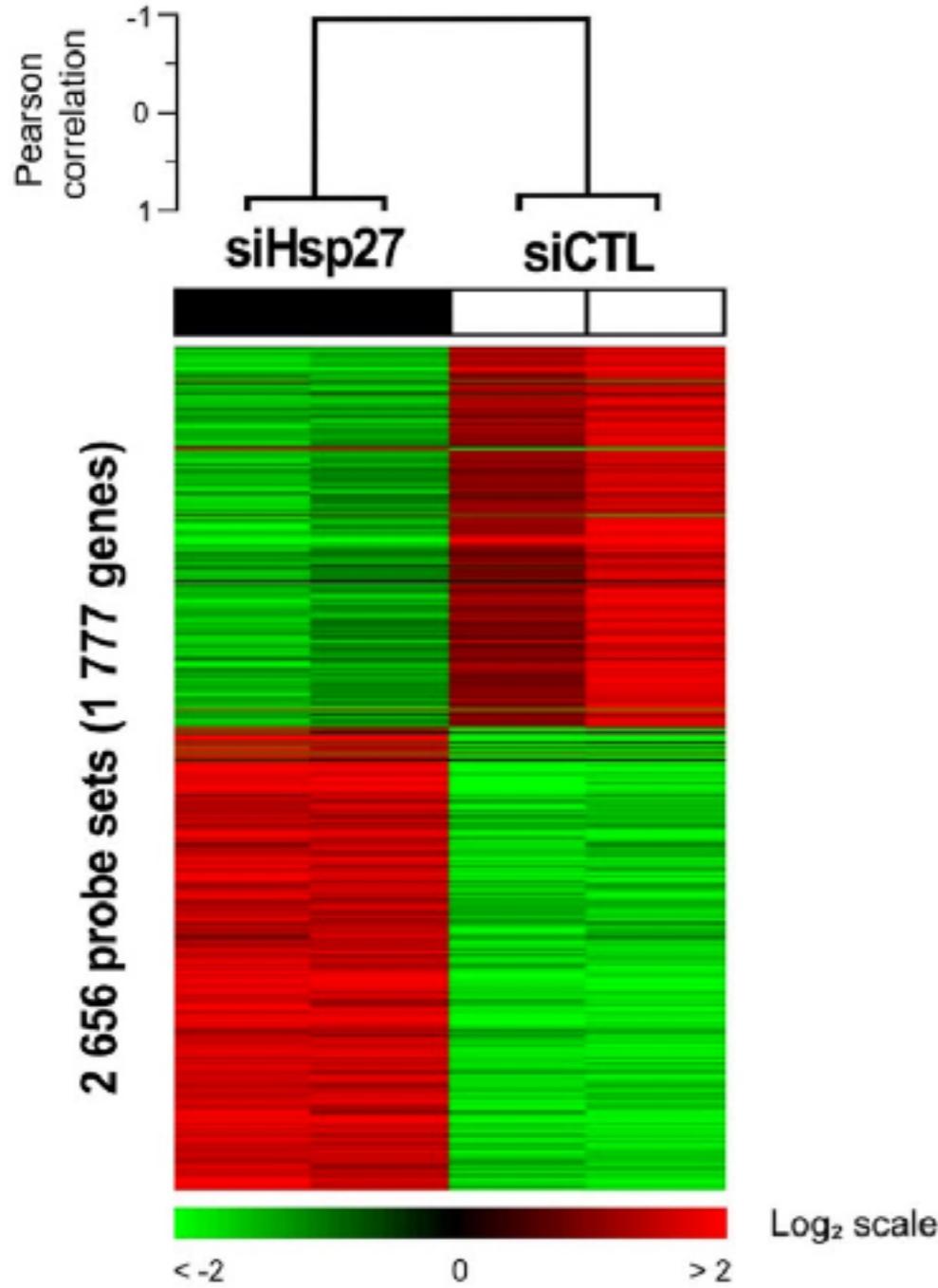


Community identification algorithms

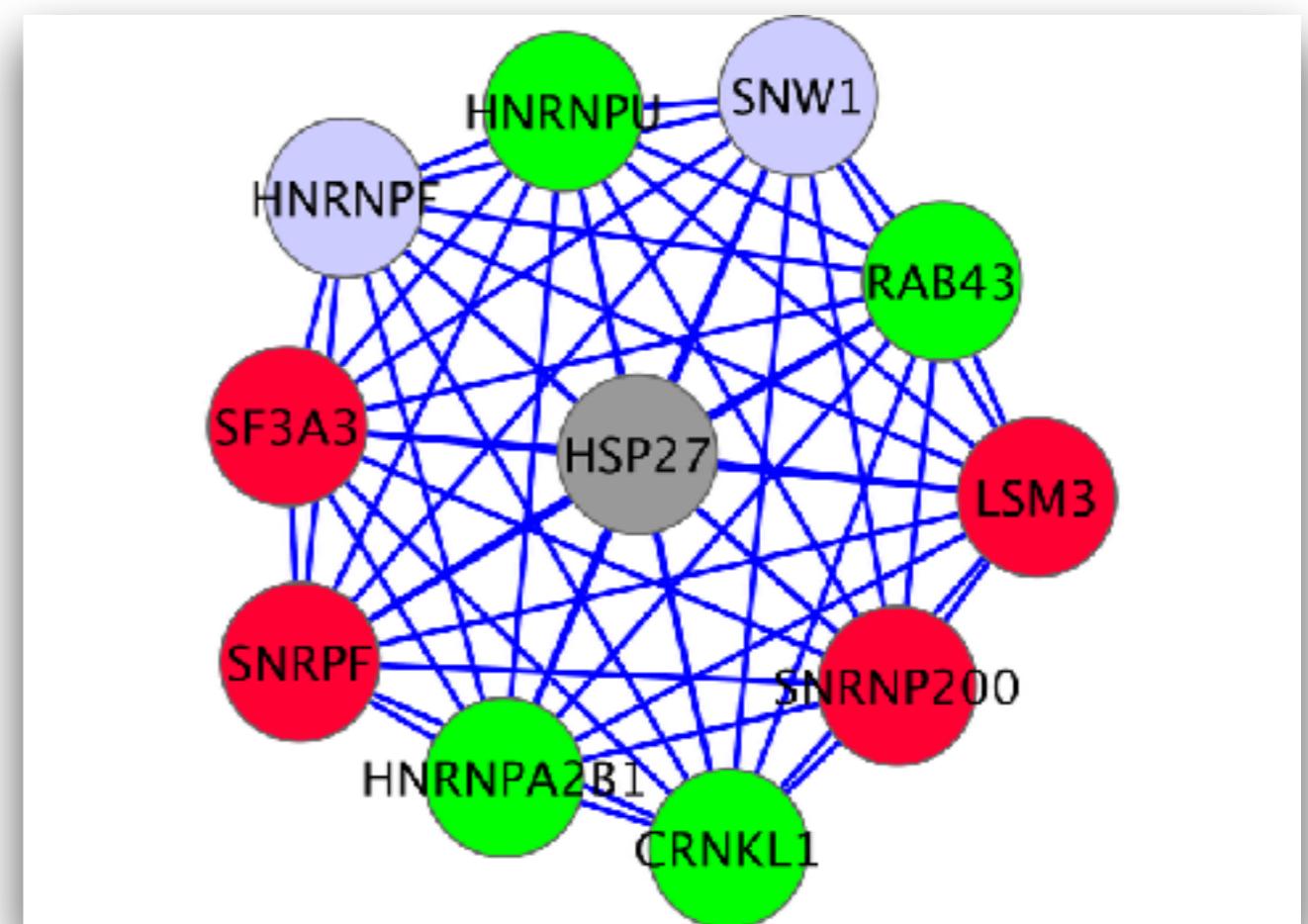


- Topology : Density, Modularity
- Overlapping / Disjoint





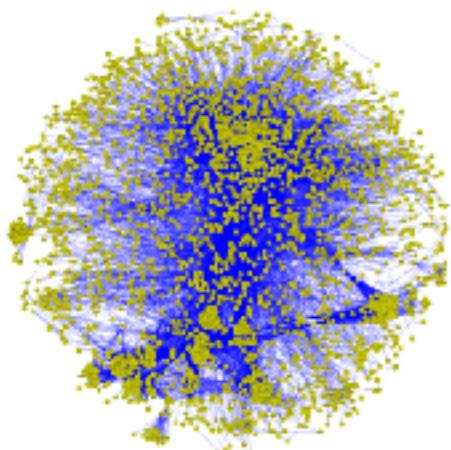
Hsp27 is involved in alternative splicing





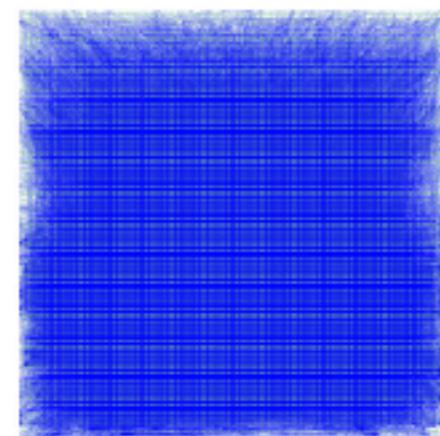
Many Biological Networks

Pathways



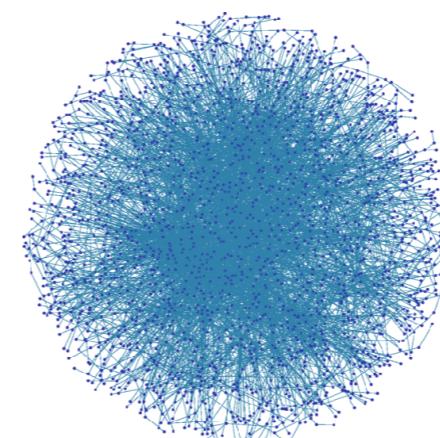
~250 000 edges

Co-expr



~1 400 000 edges

PPI



~60 000 edges

Complexes

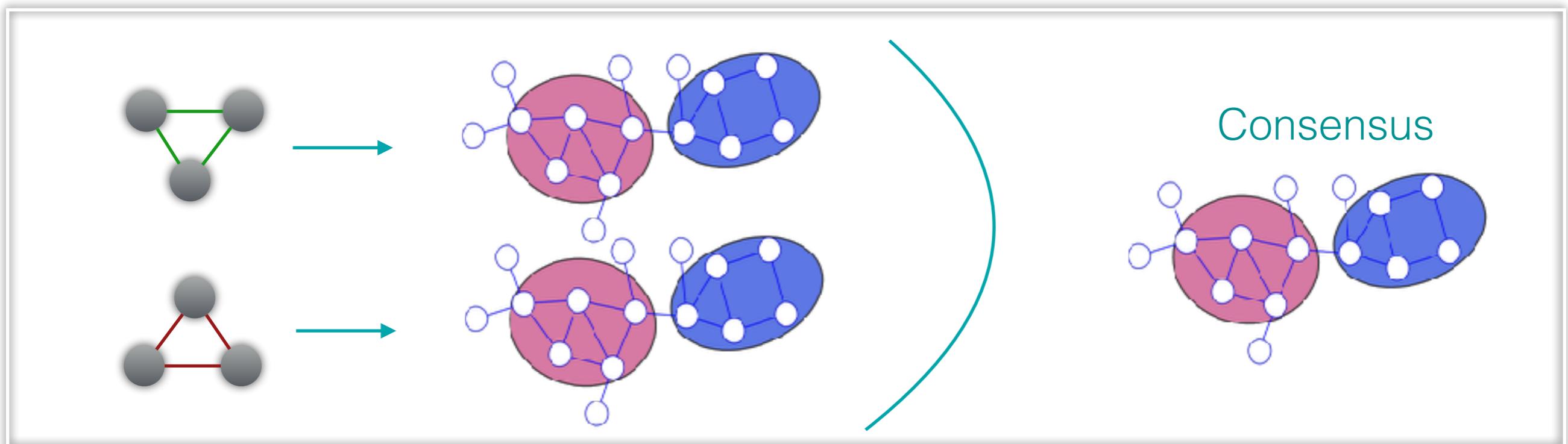
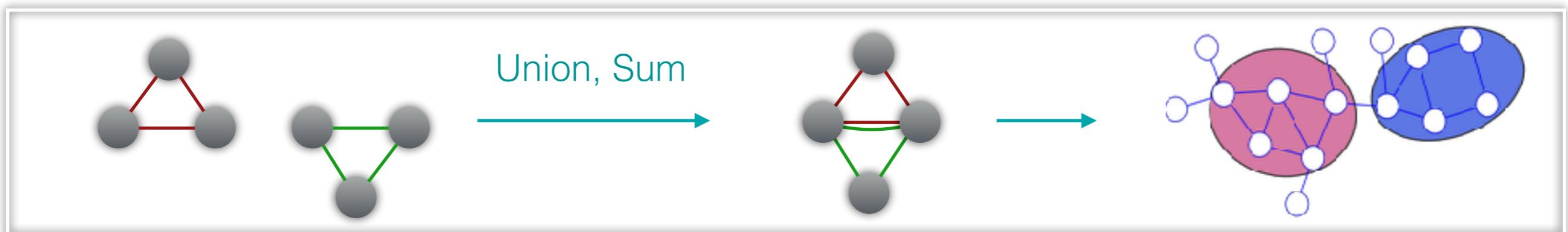


~40 000 edges

How do we combine many network/interaction sources?

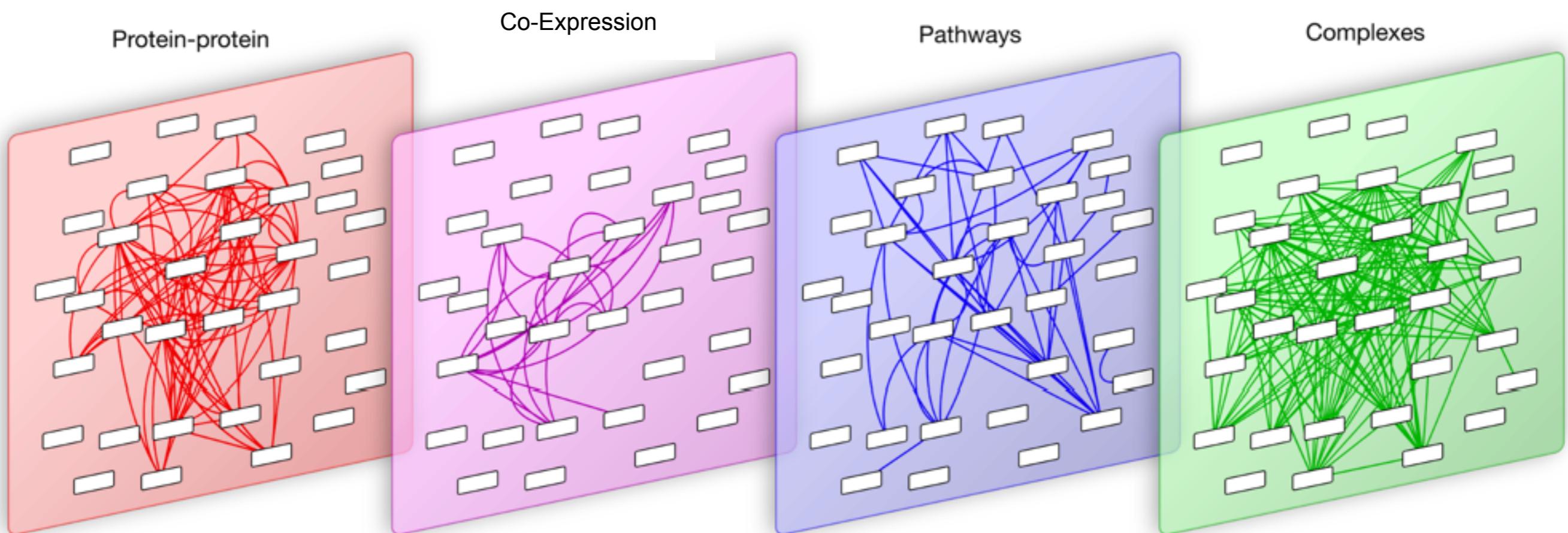


Communities from many Networks





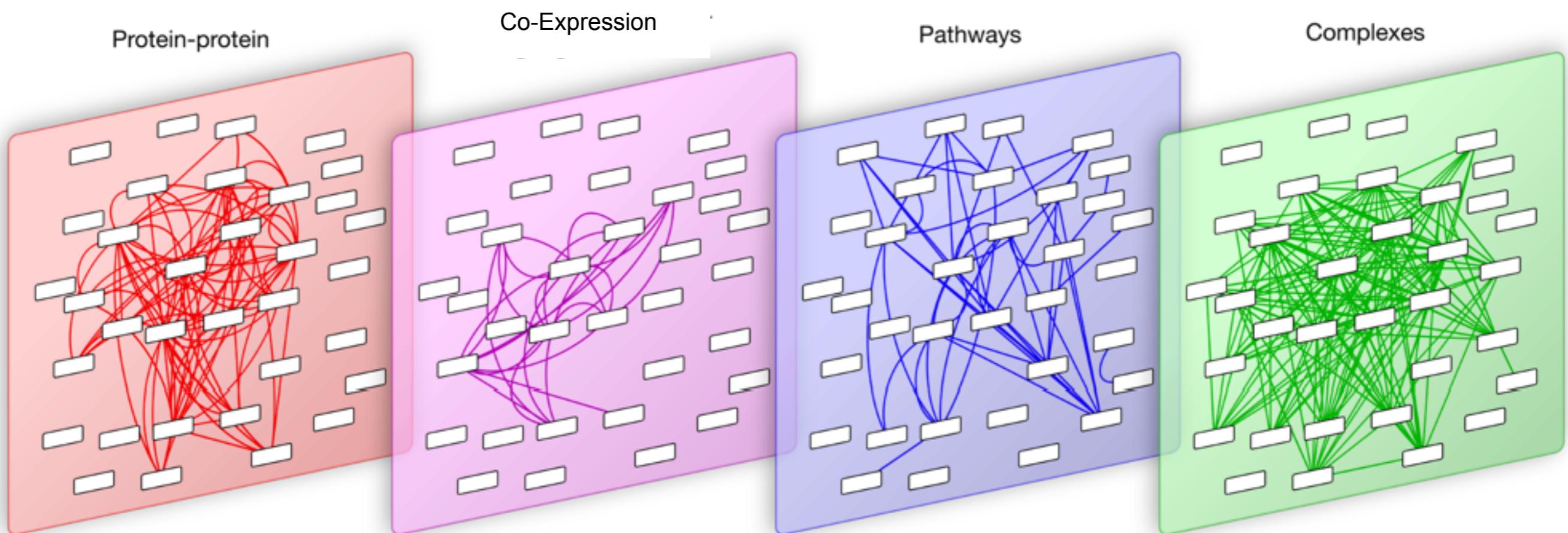
Multiplex framework



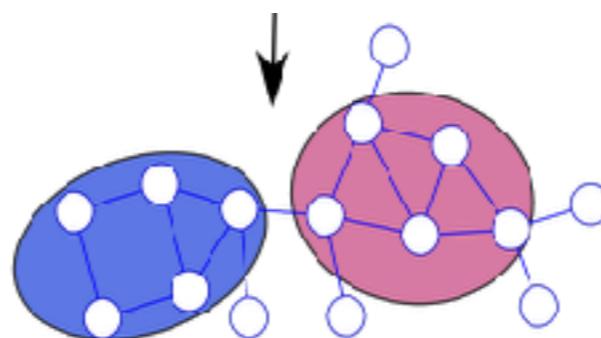
Multiplex, Multi-layer, Multislice



Multiplex framework



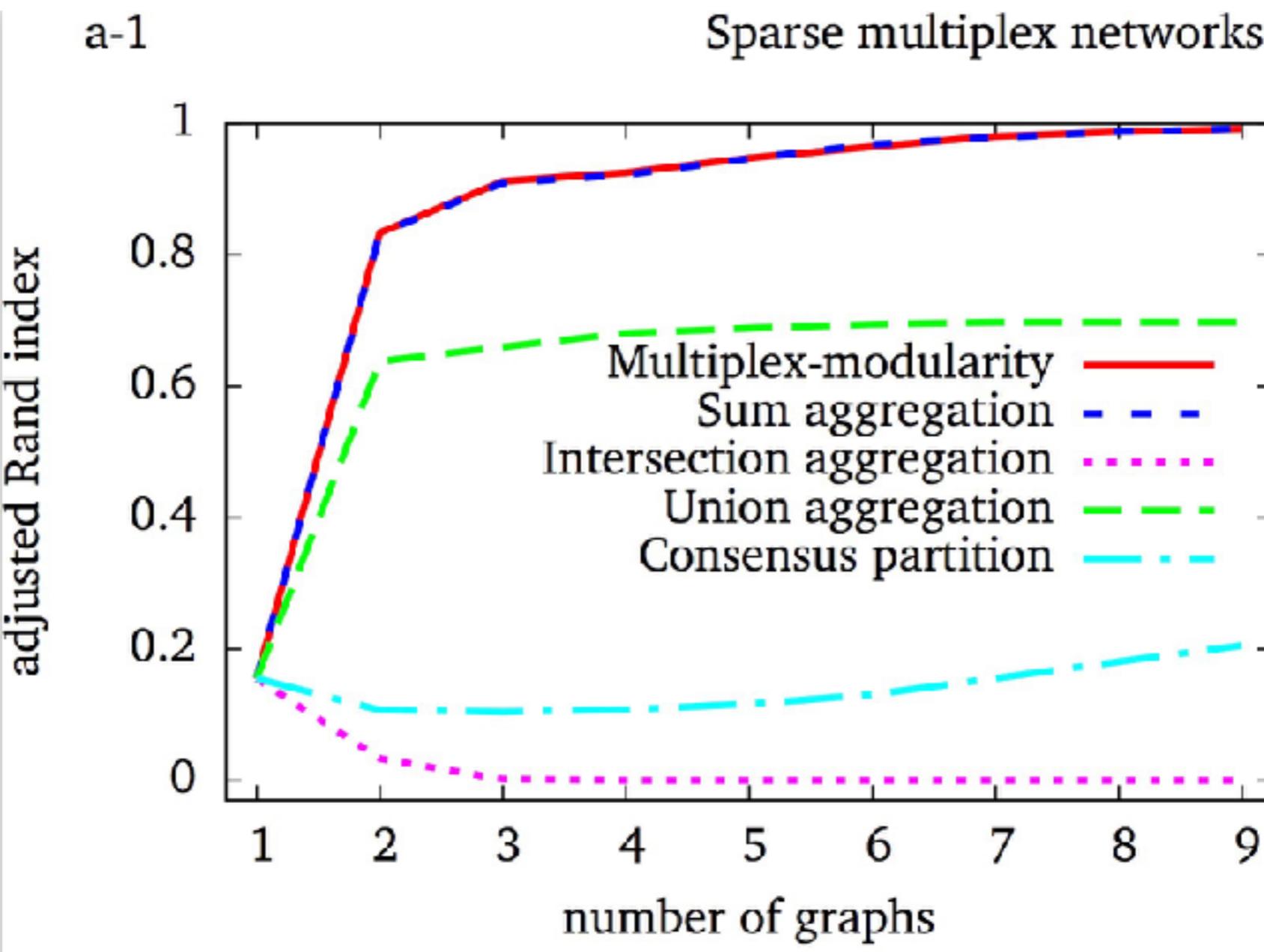
- Multiplex-modularity**



$$Q^M((X^{(g)})_g, \epsilon) = \sum_g \left[\sum_s \frac{\sum_{\{i,j\}} X_{i,j}^{(g)}}{2m^g} - \sum_g \frac{\sum_{\{i,j\}} k_i^g k_j^g}{(2m^g)^2} \right]$$



Simulations (SBMs)



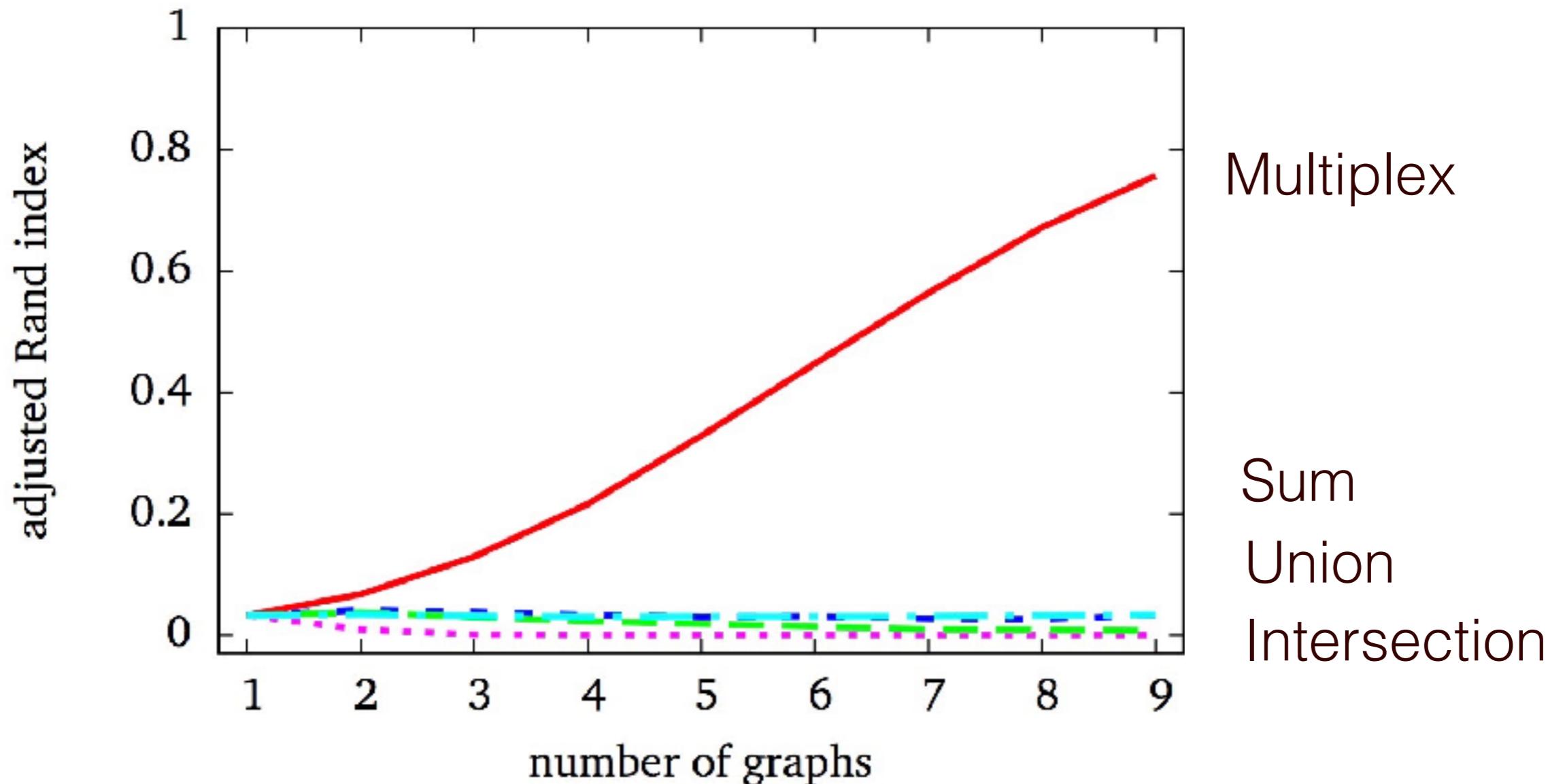
Considering many networks improves the detection of communities



Simulations (SBMs)

b-2

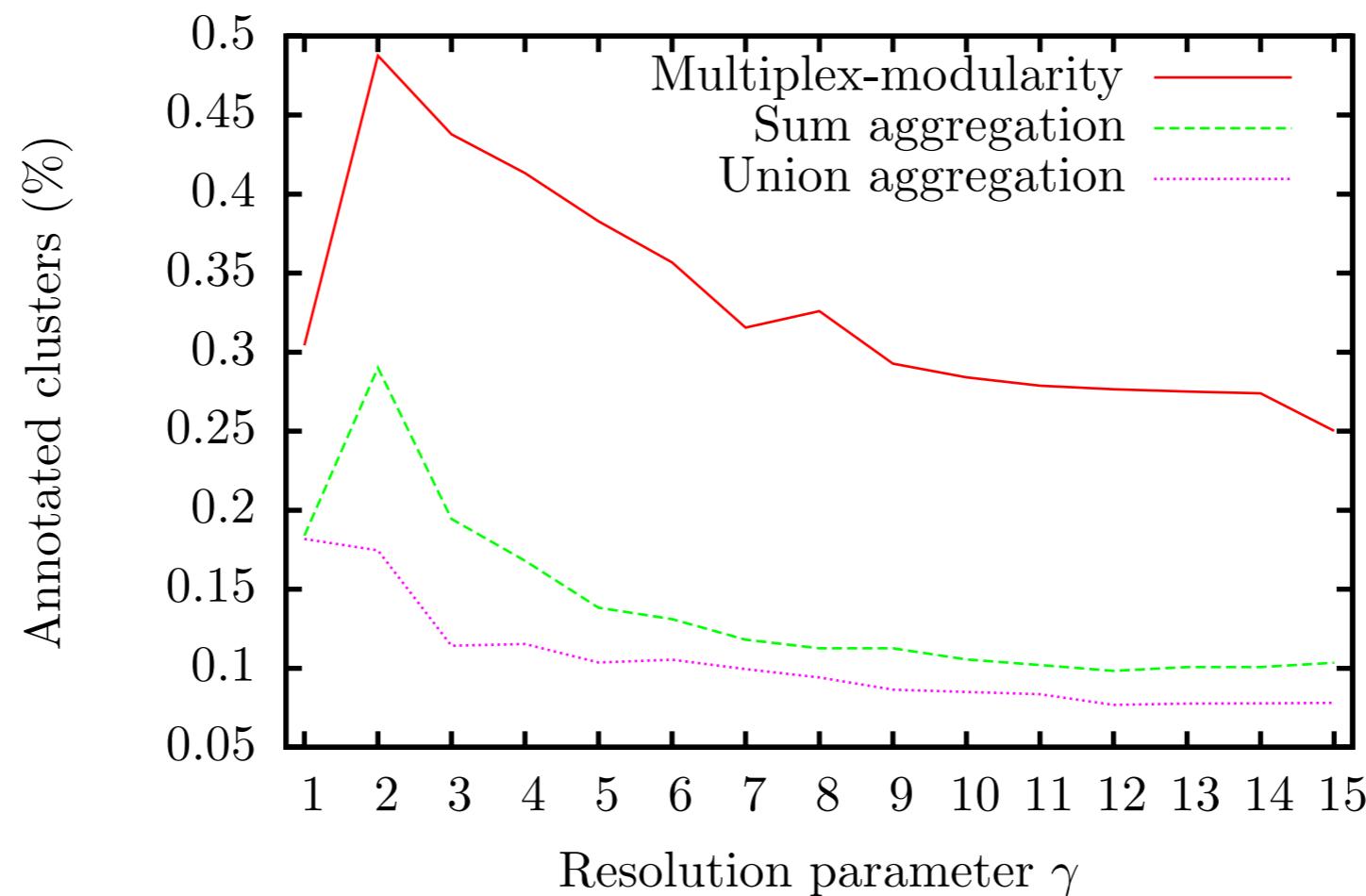
Mixed multiplex networks with missing data



Mx framework more efficient with mixed and incomplete networks



Community annotations



=> More annotated

=> Multiplex framework is well suited to identify communities from multiple biological networks



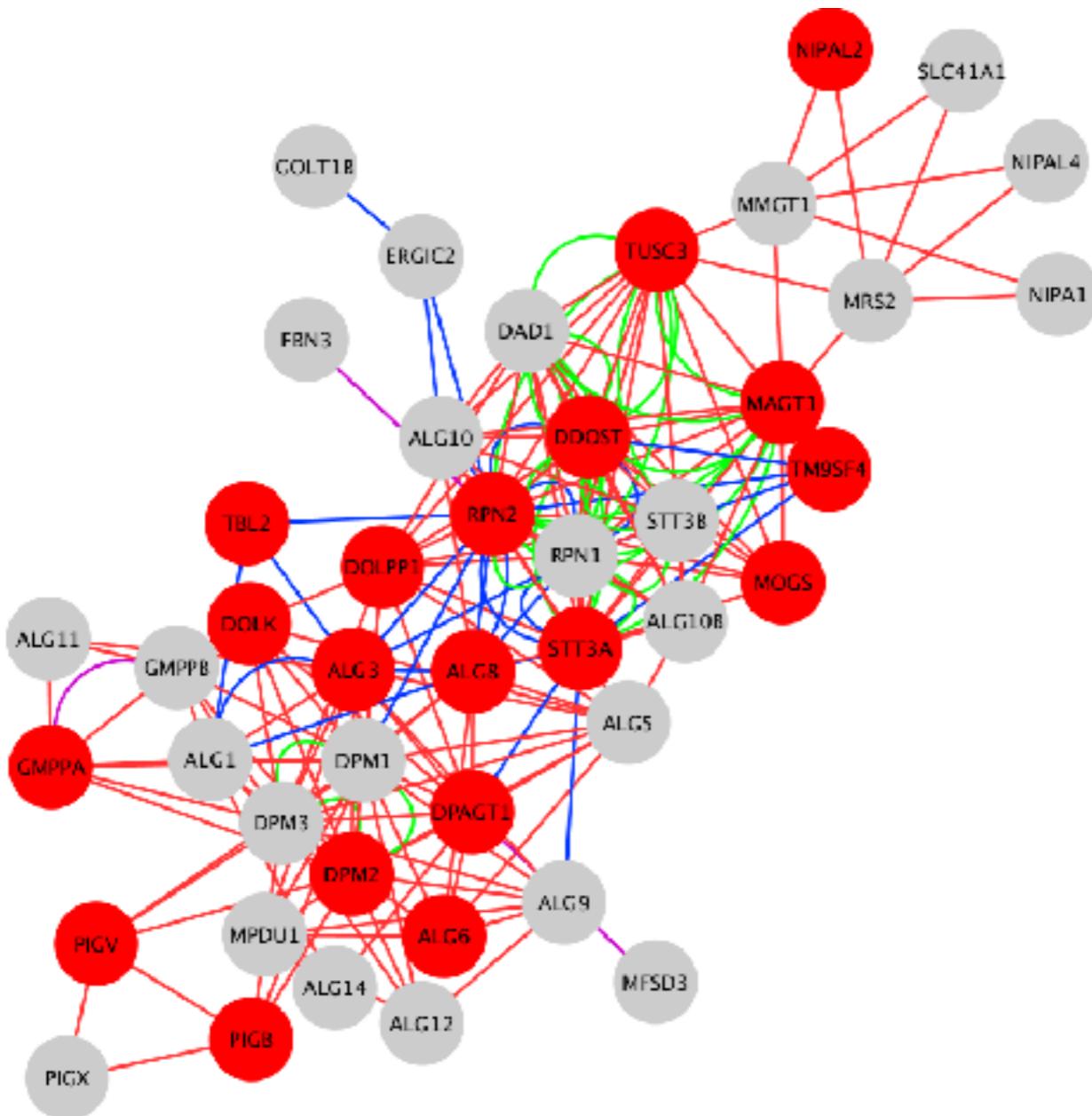
Disease-Community data integration

Find communities associated to diseases:

- Processes deregulated / impacted in diseases
- Predict new disease-associated genes
- Study disease-disease relationships



Prostate cancer up-regulated genes

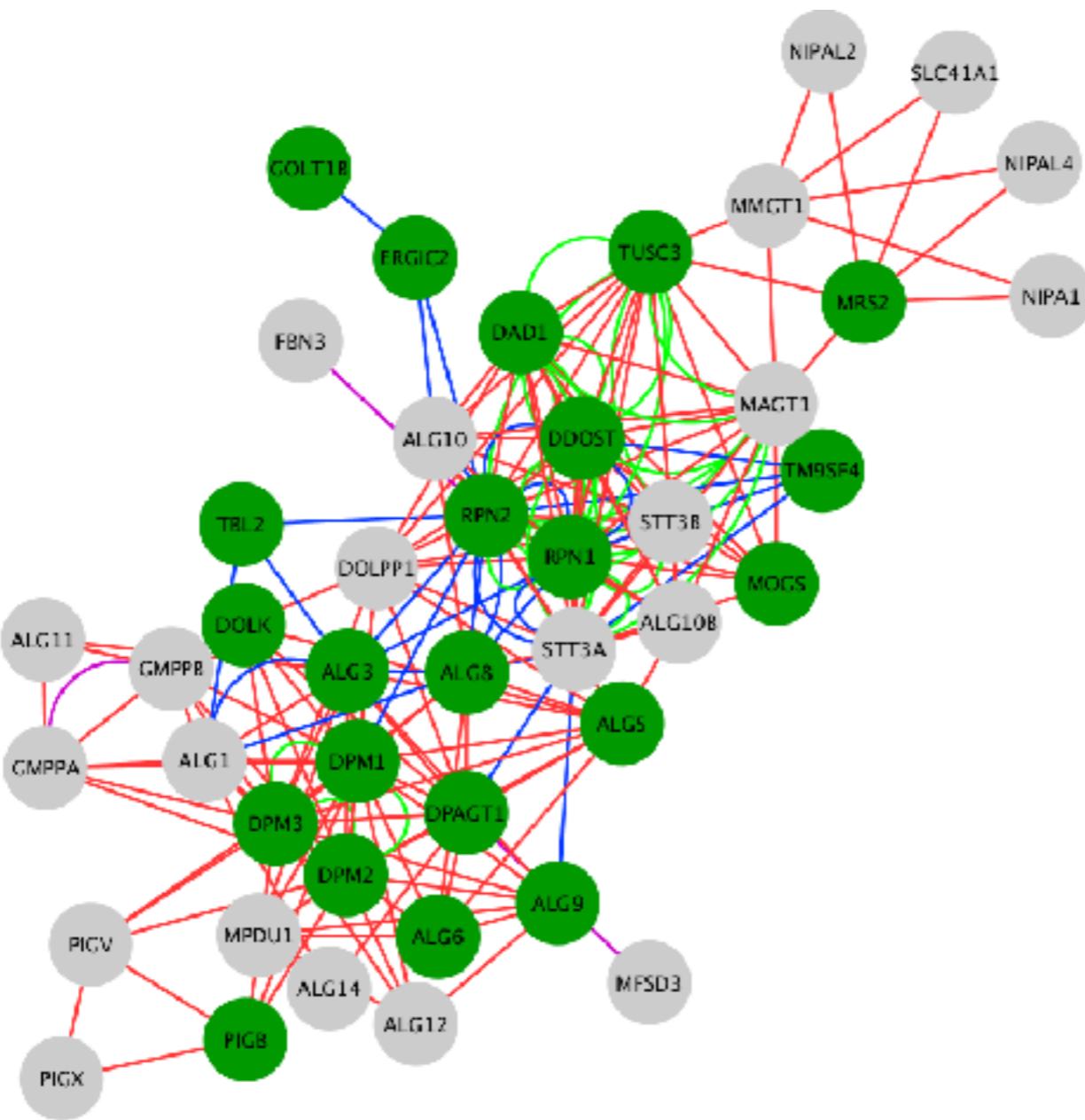


Gene expression
deregulations

glycosylation



Alzheimer's disease down-regulated genes



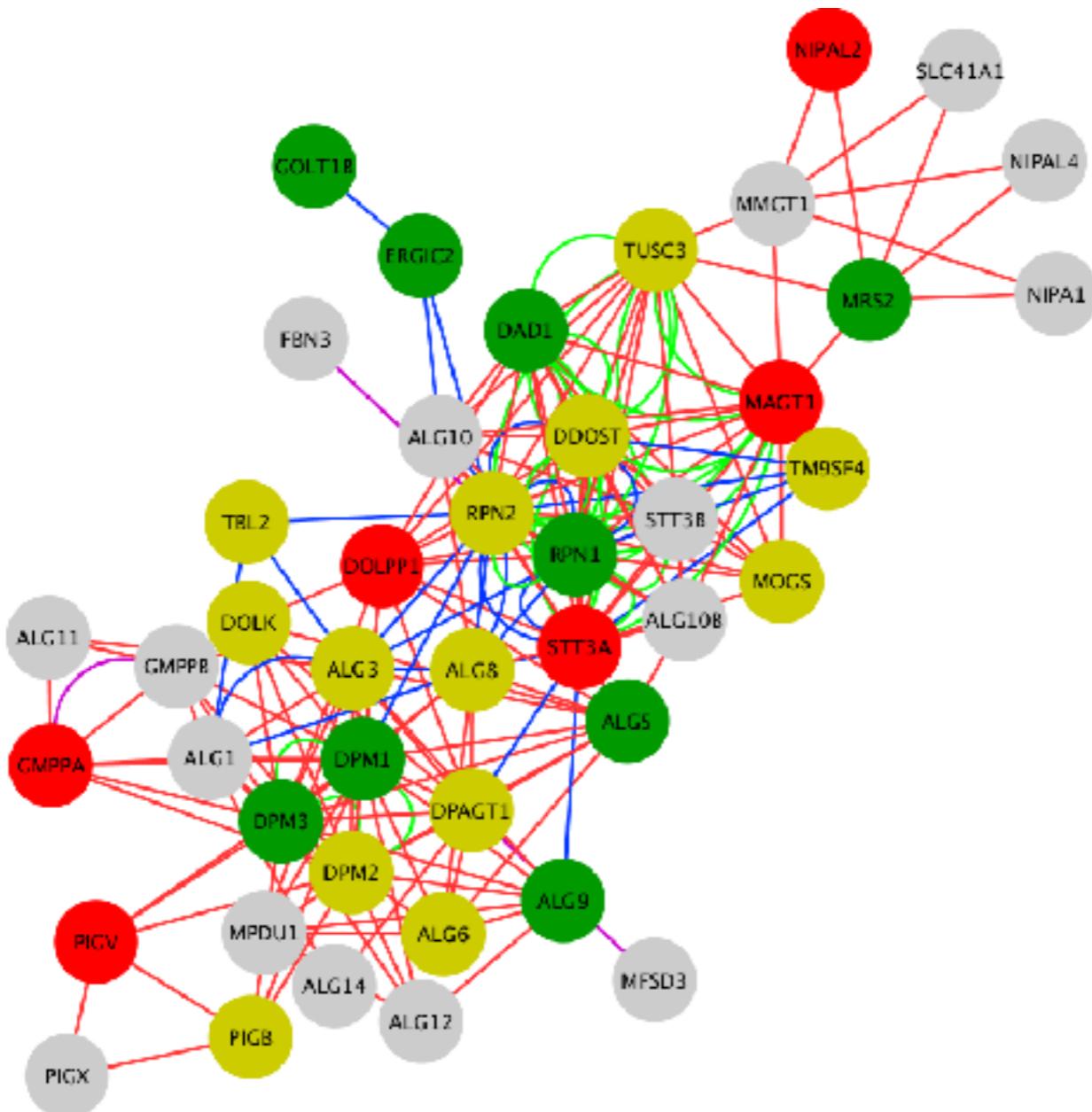
Gene expression
deregulations

glycosylation



AD & PC

- PC up
- AD down
- Both



Gene expression
deregulations

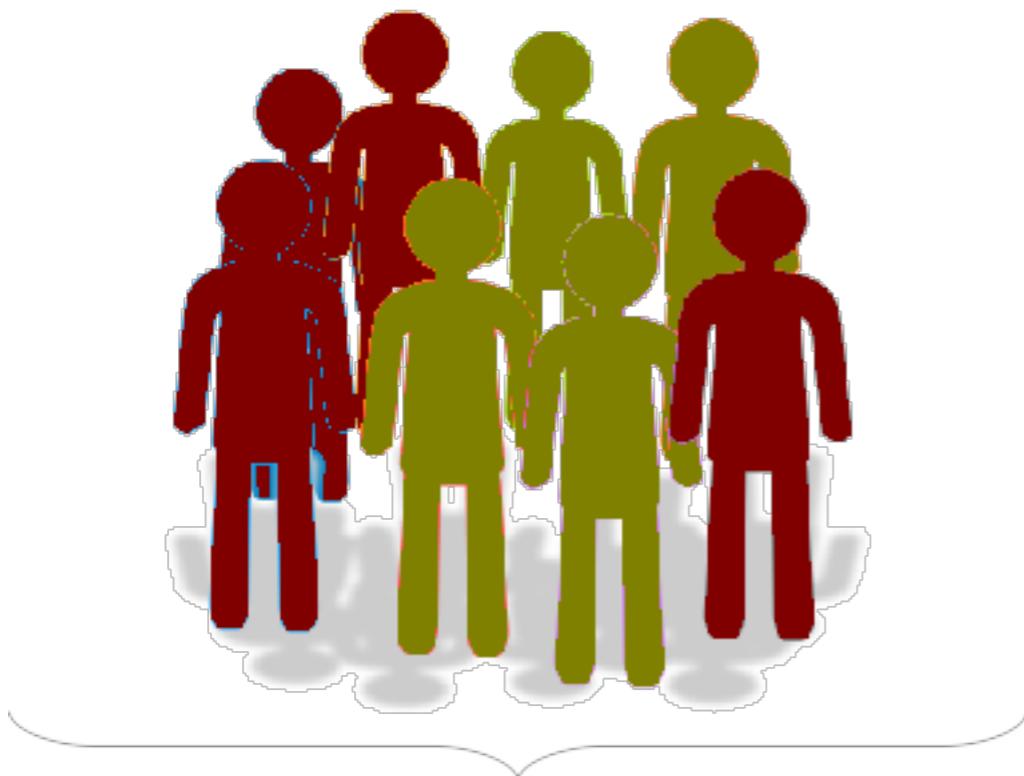
glycosylation



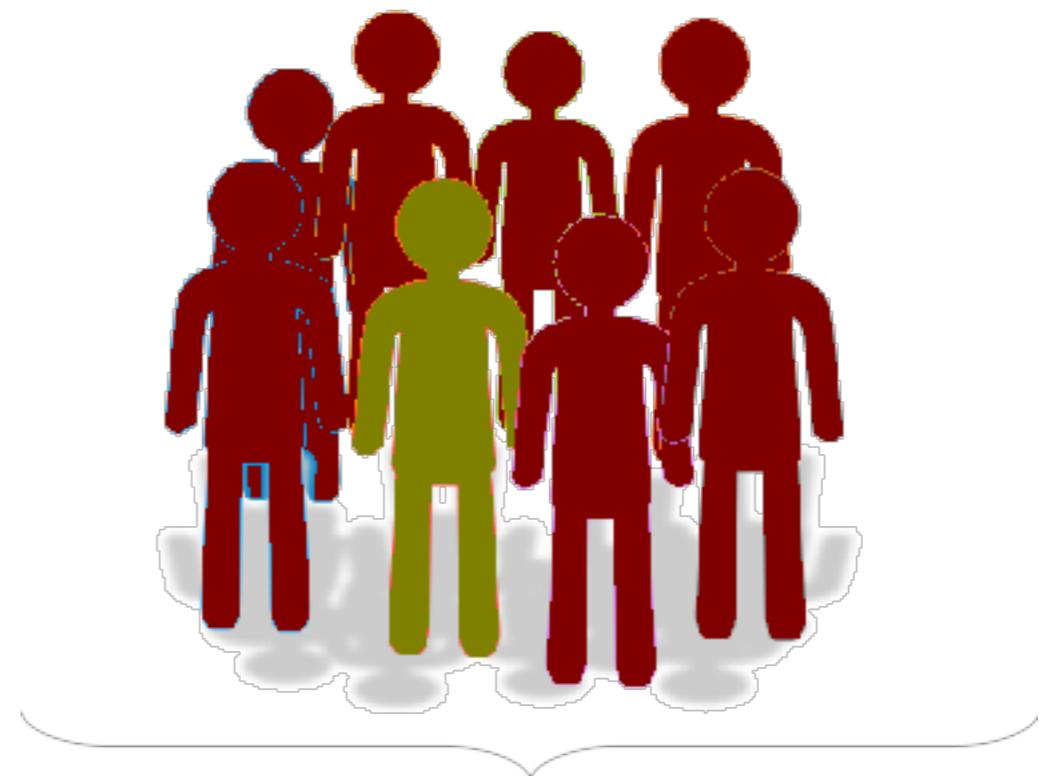
Disease **inverse comorbidity**

Inverse comorbidity / Multimorbidity
Decreased co-occurrence of diseases

Ex: Alzheimer's disease and cancers



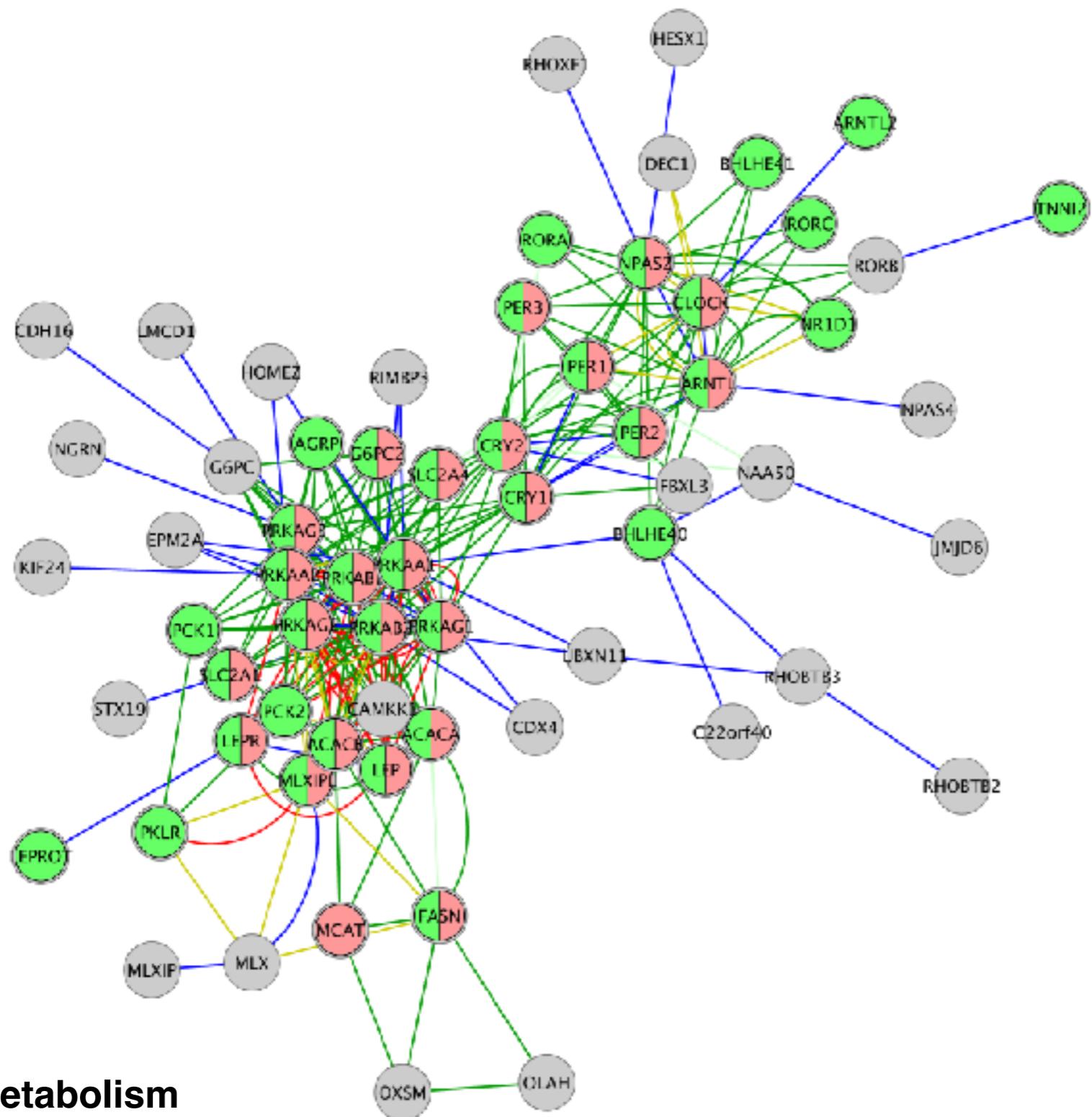
Control



Alzheimer's Disease



Susceptibility genes (GWAS data)



Prostate Cancer

Diabetes Mellitus



Fatty acid metabolism

Circadian rhythm



Next step => tissue expression integration

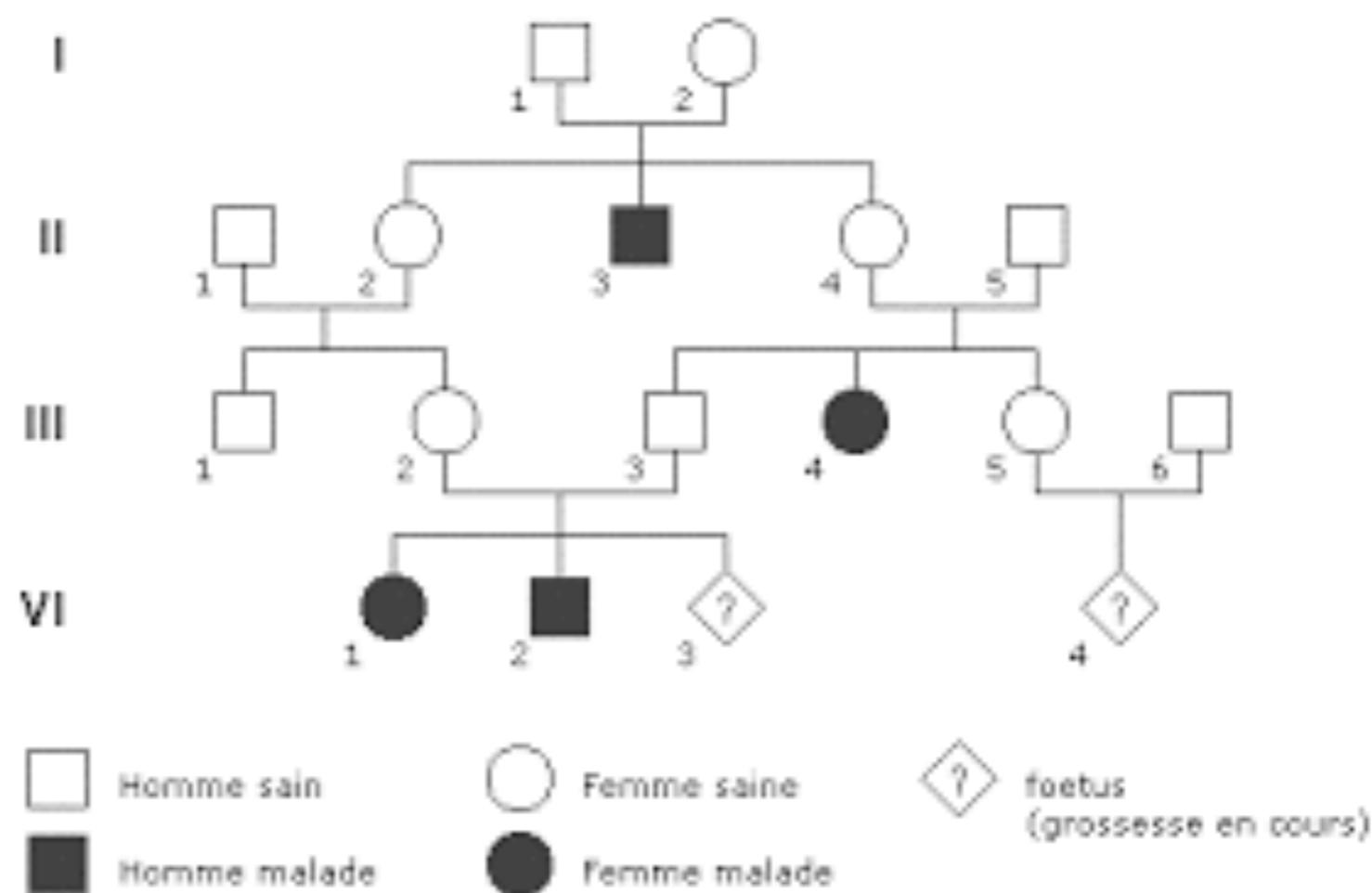


- **Transcriptomics**
- **Proteomics**



Rare monogenic diseases

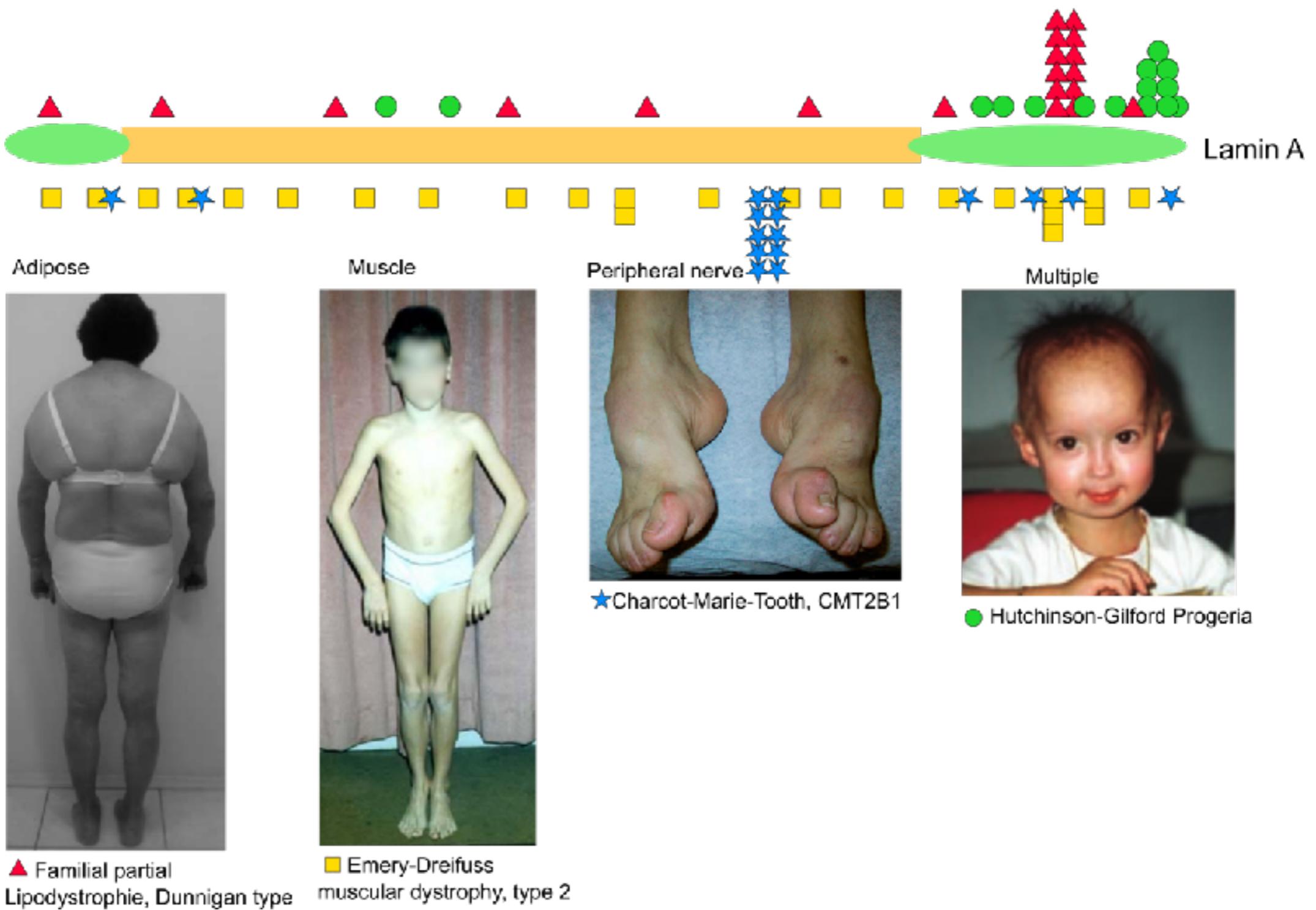
- Definitions
 - Also very complex ...



Clinical heterogeneity



Ex: Laminopathies





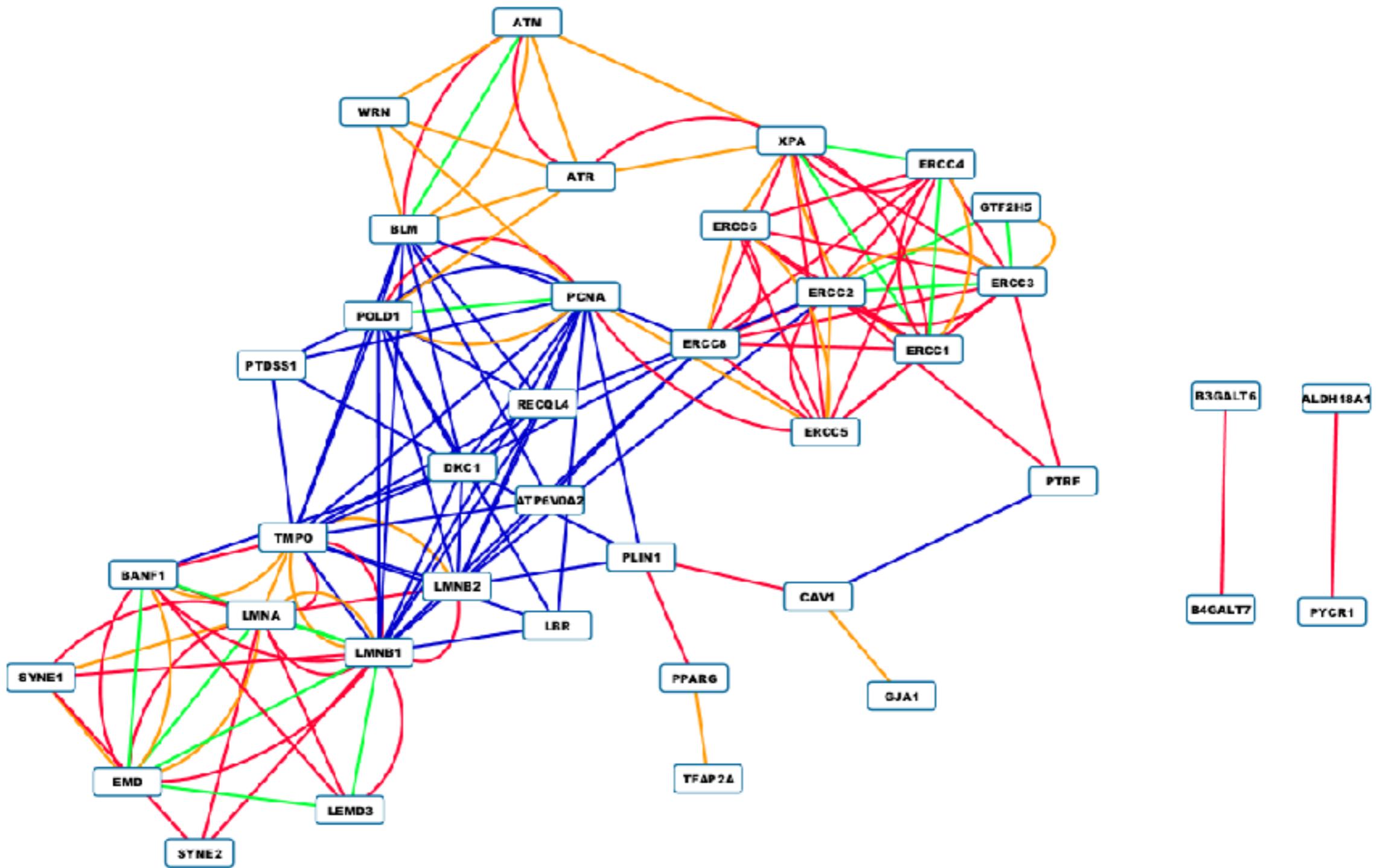
Locus heterogeneity

Ex: Premature Aging Syndromes

		liste_genes_vieillissement_prematuré.txt
BANF1	NM_003860	
BLM	NM_000057	
ERCC1	NM_001983	
ERCC2	NM_000400	
ERCC3	NM_000122	
ERCC6	NM_000124	
ERCC8	NM_000082	
GTF2H5	NM_207118	
LMNA	NM_170707	
RECQL4	NM_004260	
MPLKIP	NM_138701	
WRN	NM_000553	
XPA	NM_000380	
ERCC5	NM_000123	
ZMPSTE24	NM_005857	
PIK3R1	NM_181523	
B4GALT7	NM_007255	
B3GALT6	NM_080605	
ALDH18A1	NM_002860	
ATP6V0A2	NM_012463	
ATM	NM_000051	
ATR	NM_001184	
GJA1	NM_000165	
PTDSS1	NM_014754	
ERCC4	NM_005236	
POLD1	NM_002691	
FBN1	NM_000138	
TFAP2A	NM_003220	
PCNA	NM_002592	
TERC	NR_001566	
DKC1	NM_001363	
PYCR1	NM_006907	

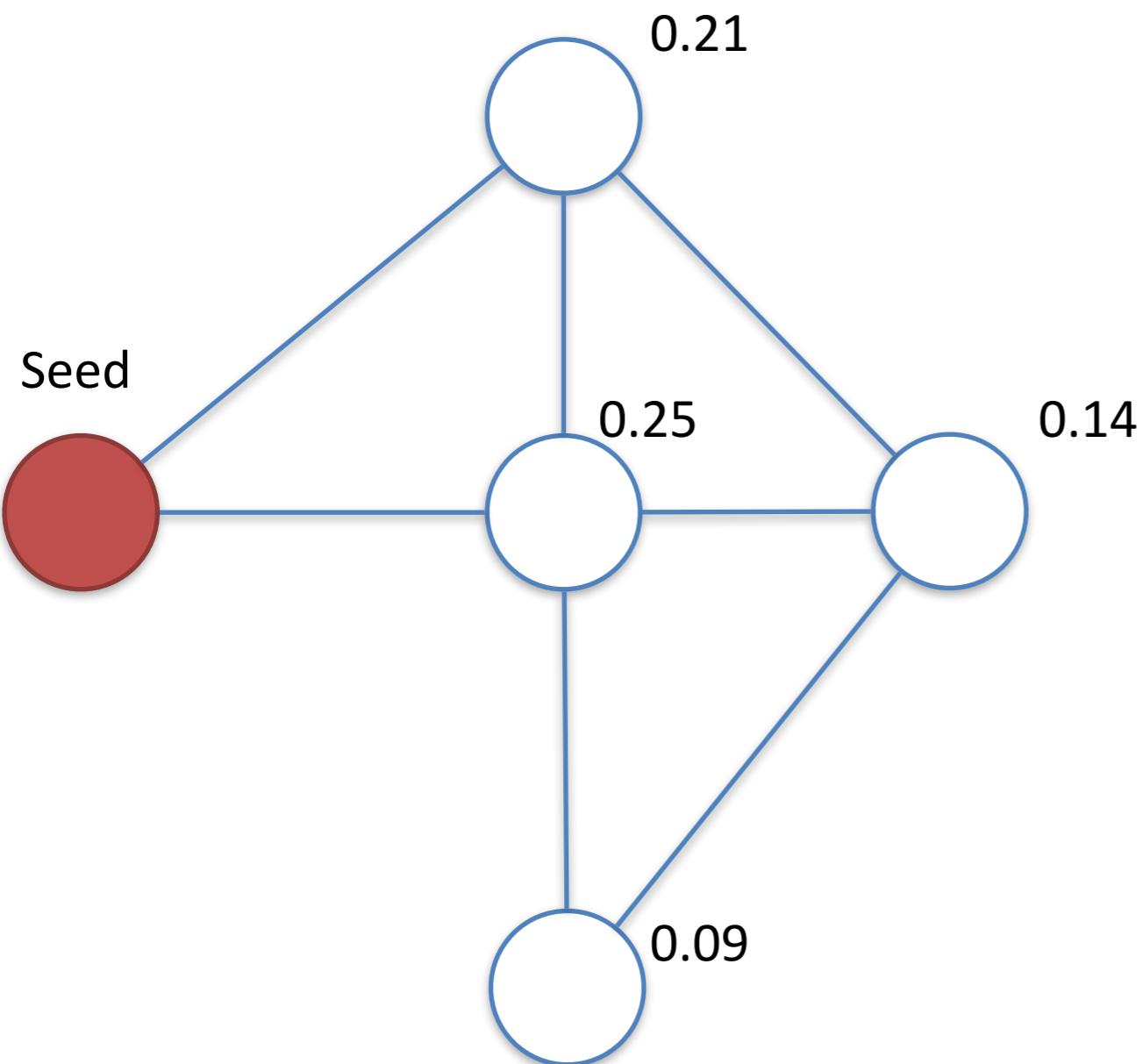


Premature aging interactome



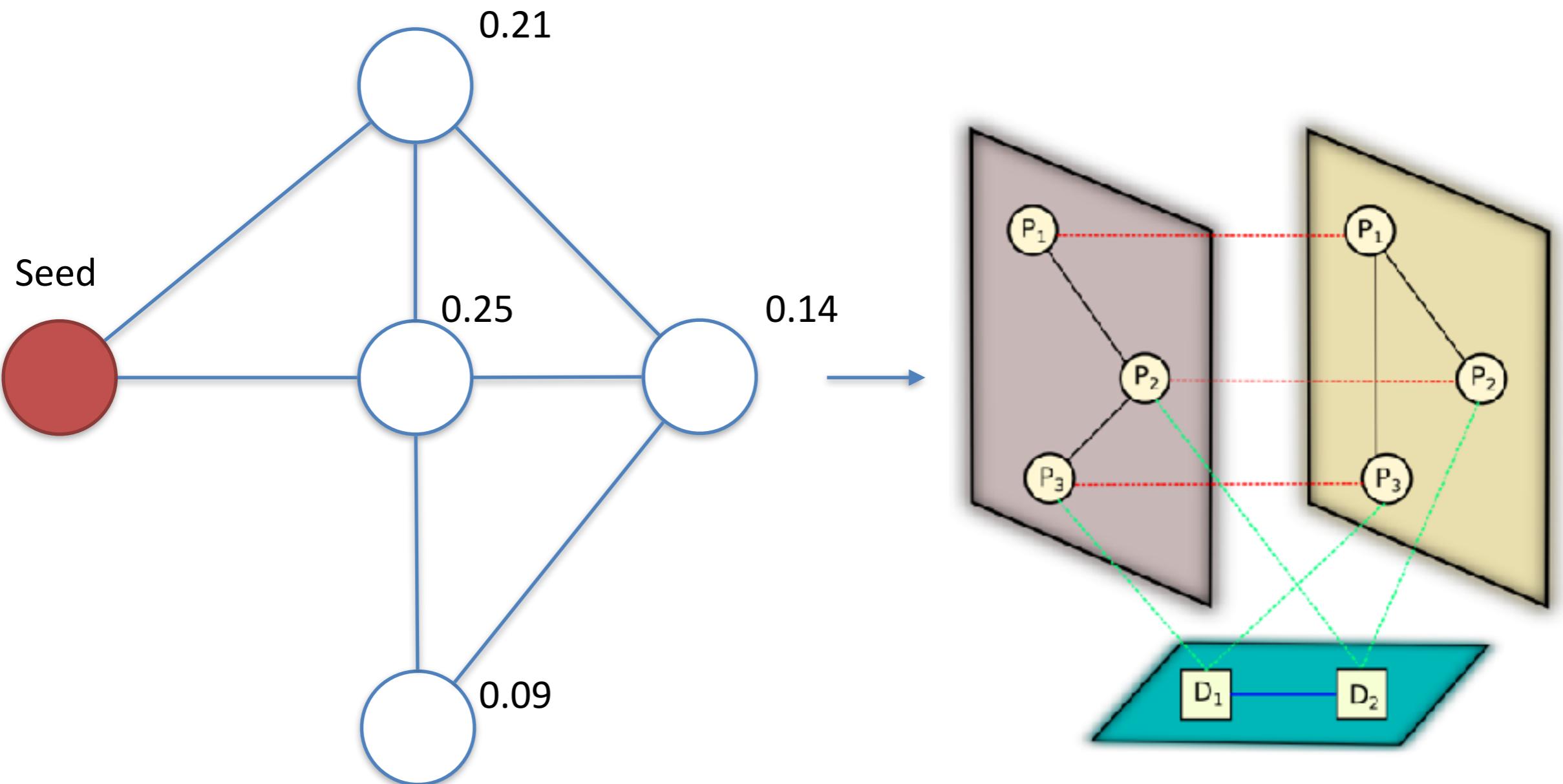


Random Walk with Restart





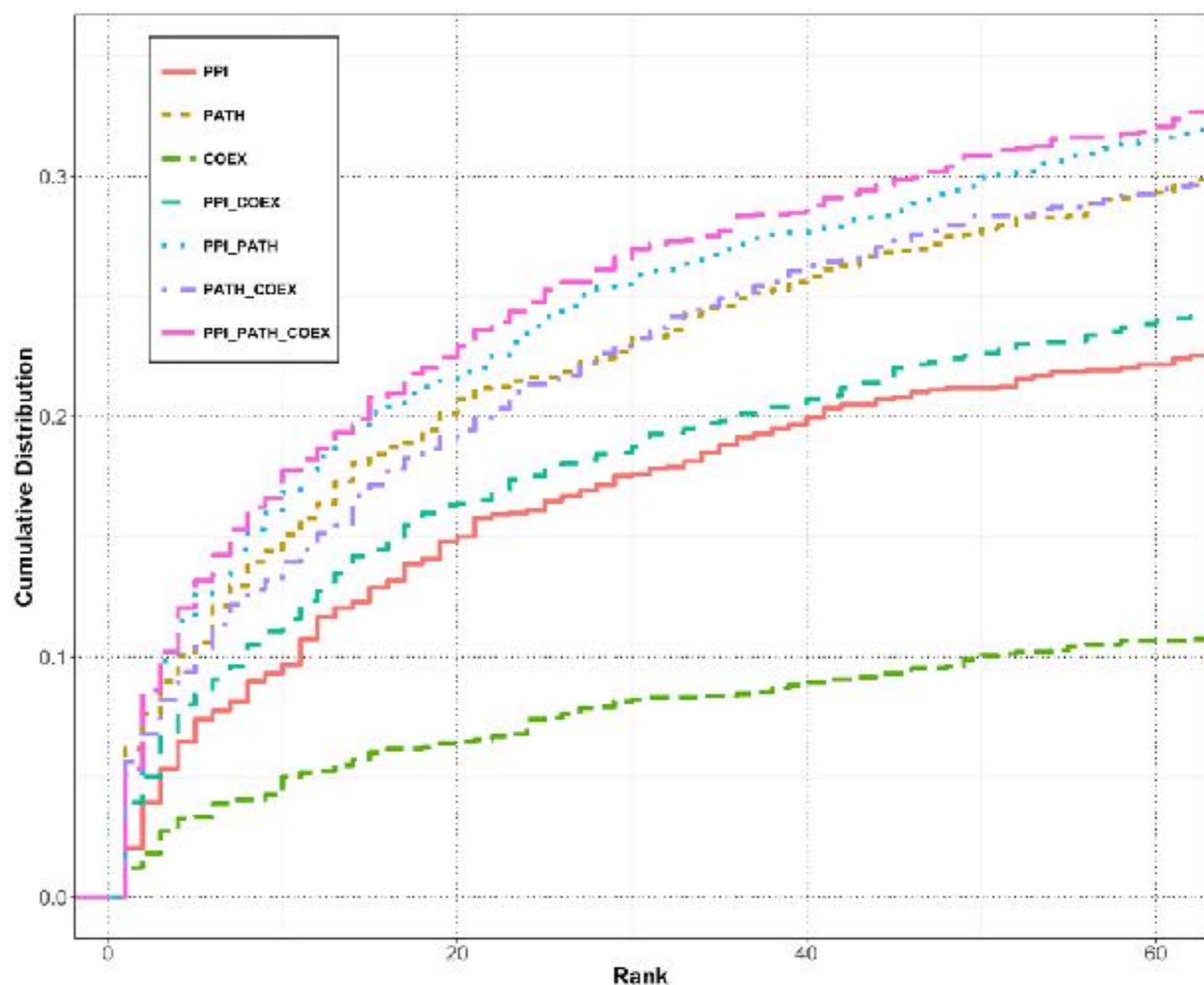
Random Walk with Restart in Multiplex and Heterogeneous Networks





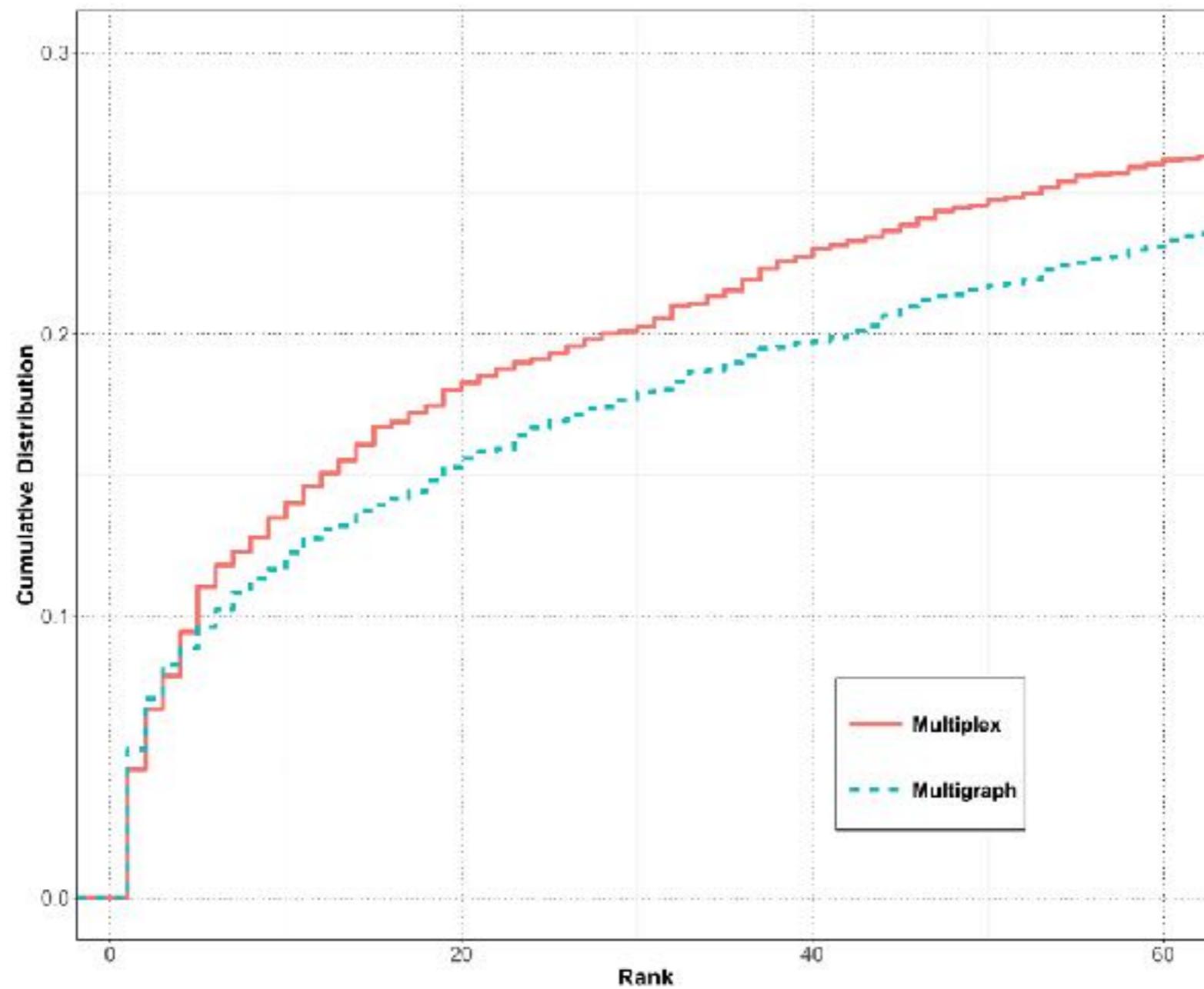
RWR on Multiplex versus Monoplex

Leave-One-Out Cross Validation

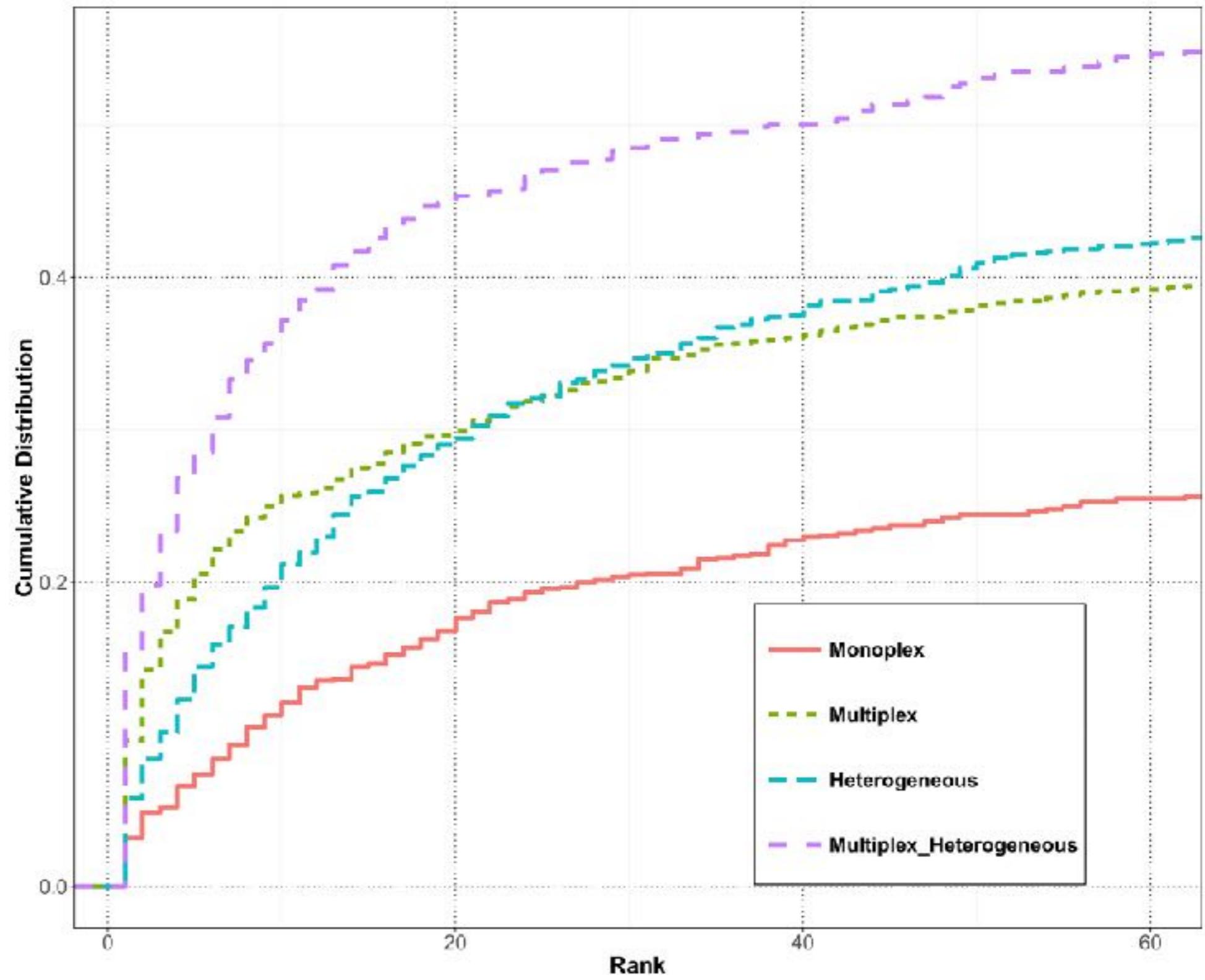




RWR on Multiplex versus aggregated network

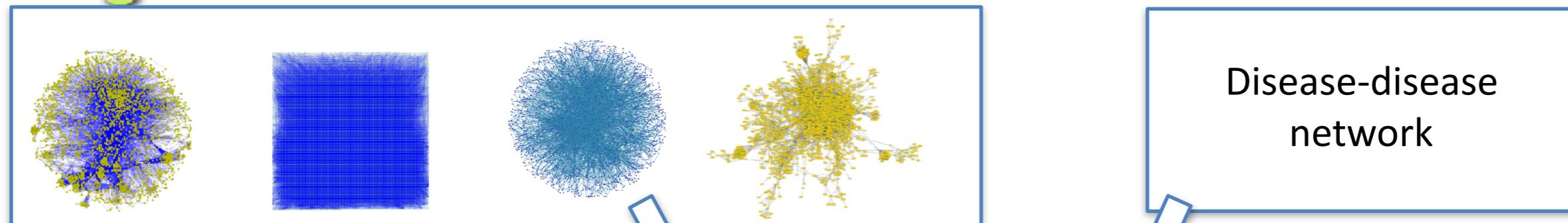


RWR on Multiplex and/or Heterogeneous network

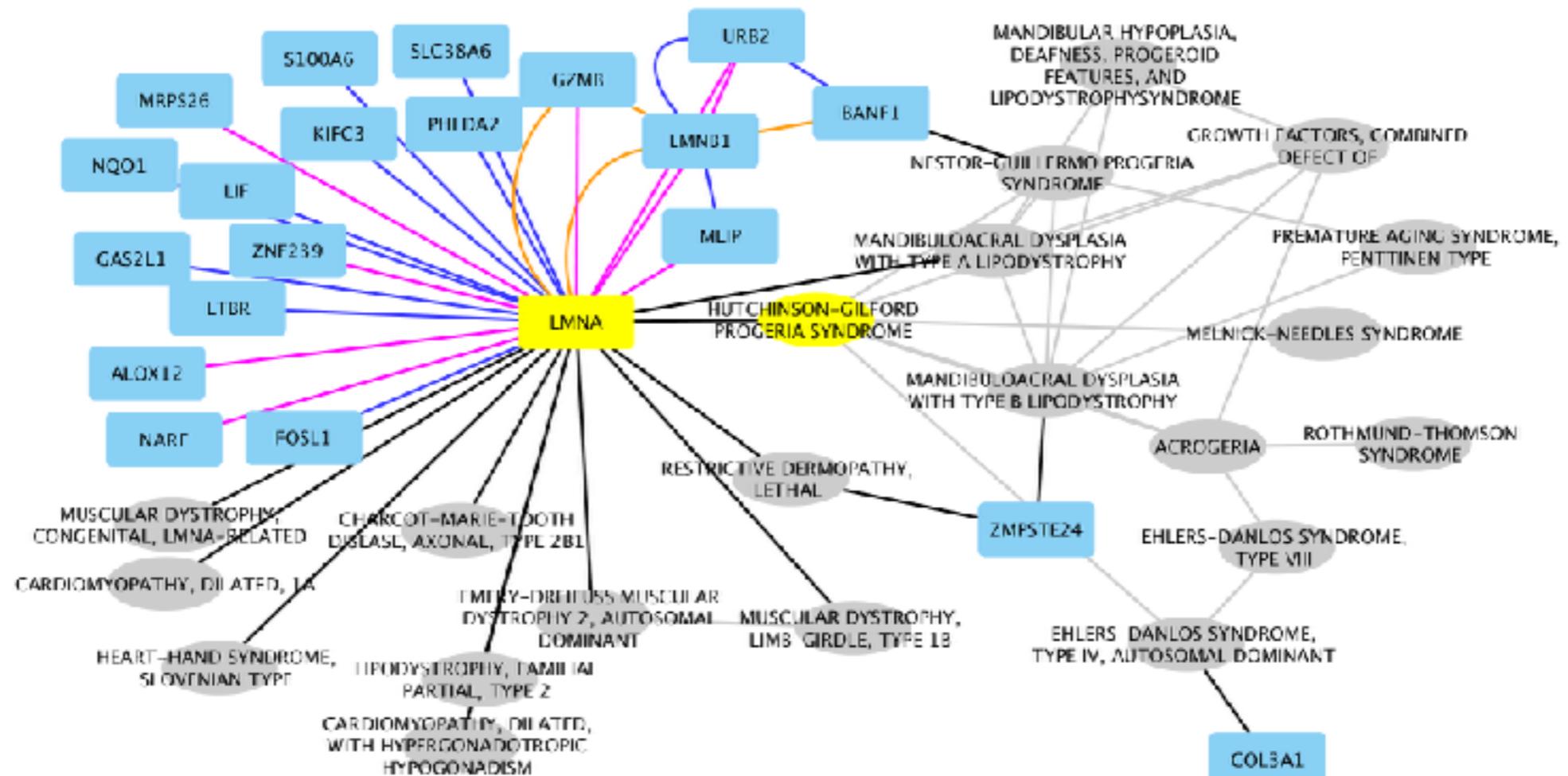




Application: Explore Network vicinity of disease and disease gene



Random walk
with restart





Application: From premature to physiological aging

Rare diseases

Common phenotypes



Use rare premature diseases as “seeds” to:

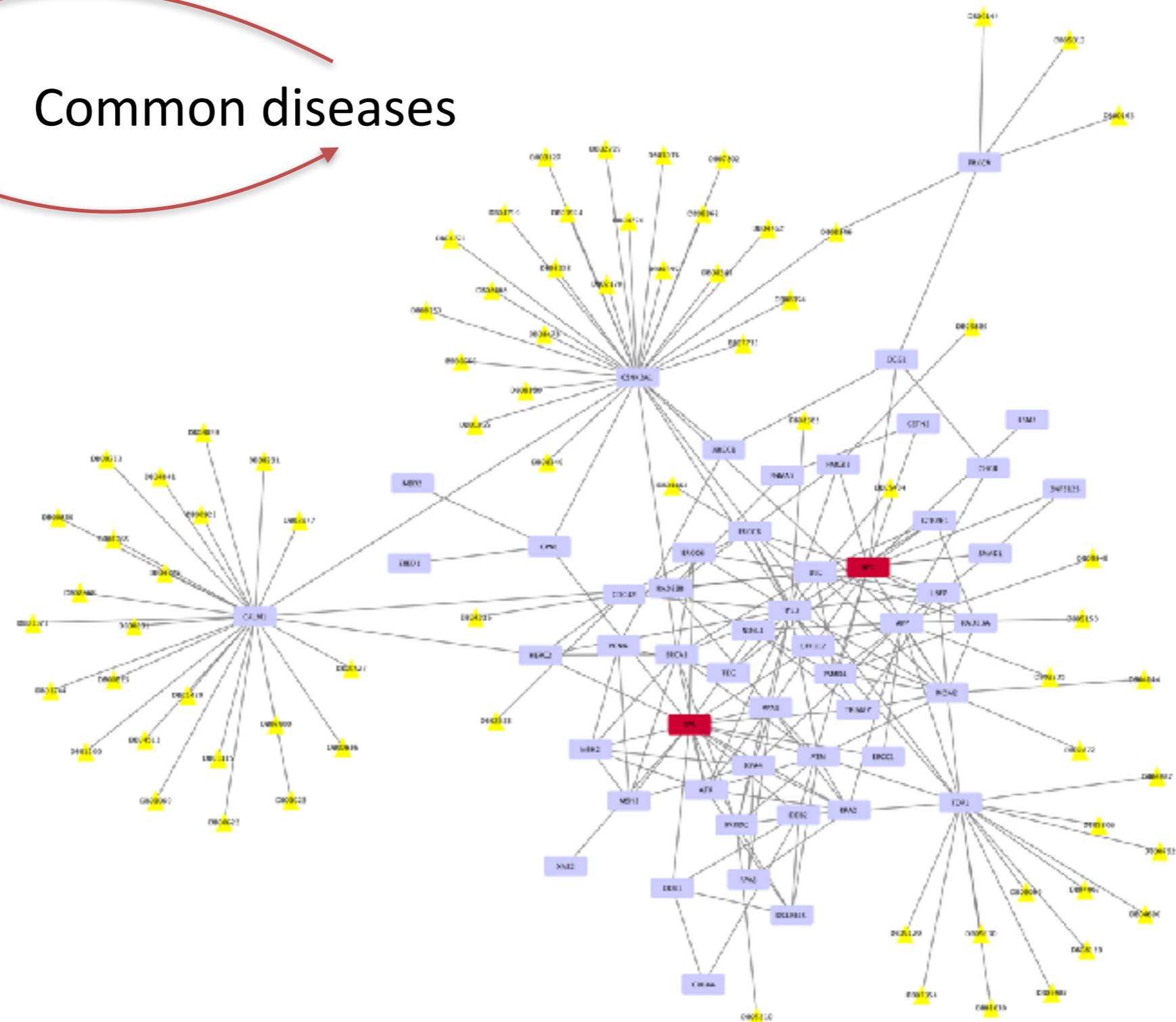
- define the molecular landscape of aging
- find new biomarkers of aging

Application for drug repurposing

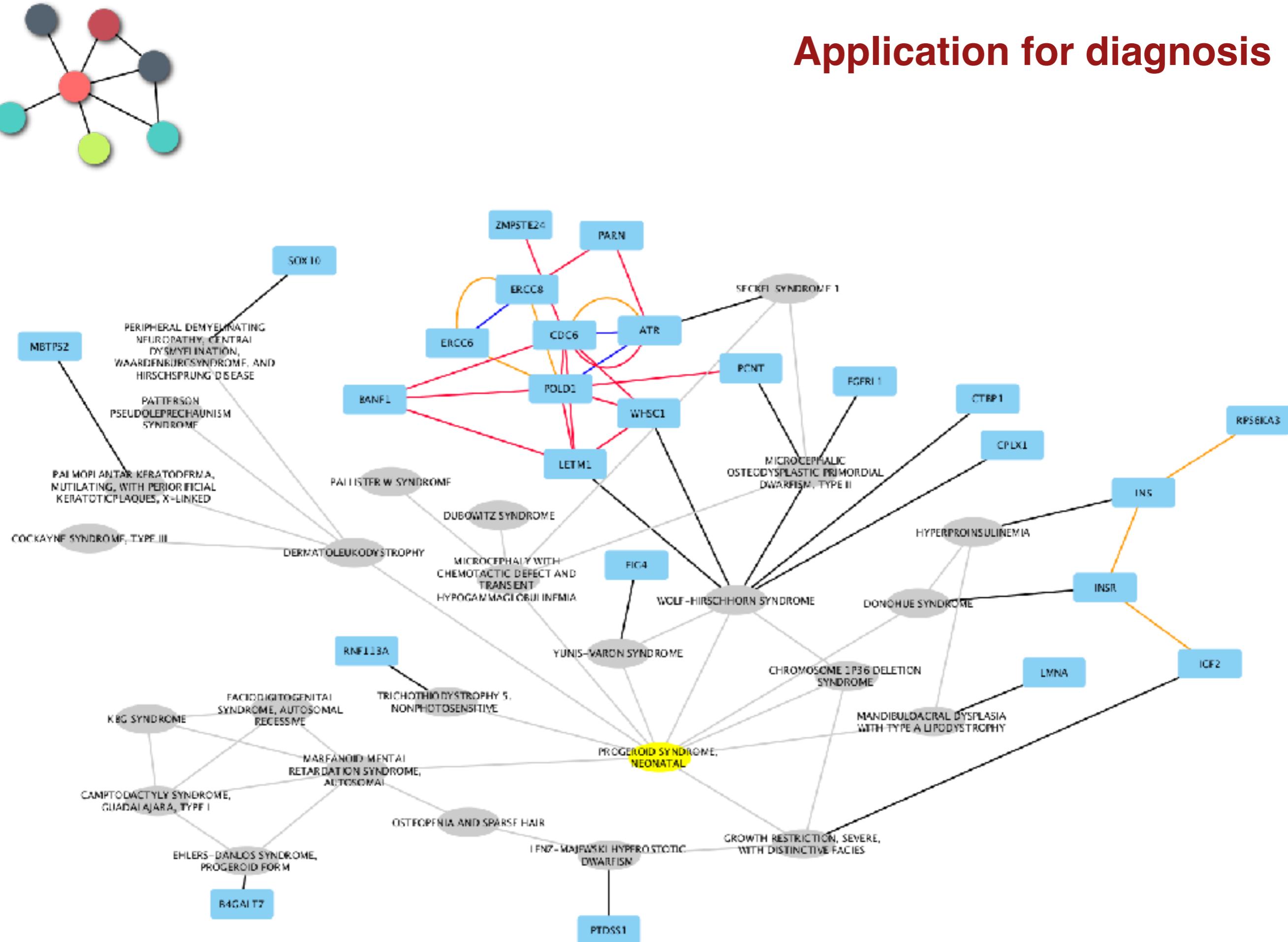


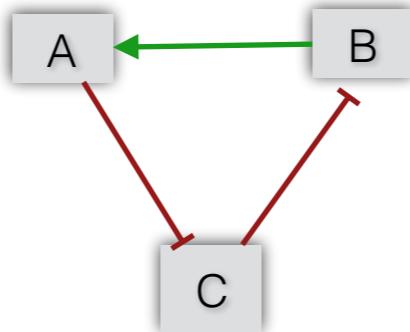
Rare diseases

Common diseases



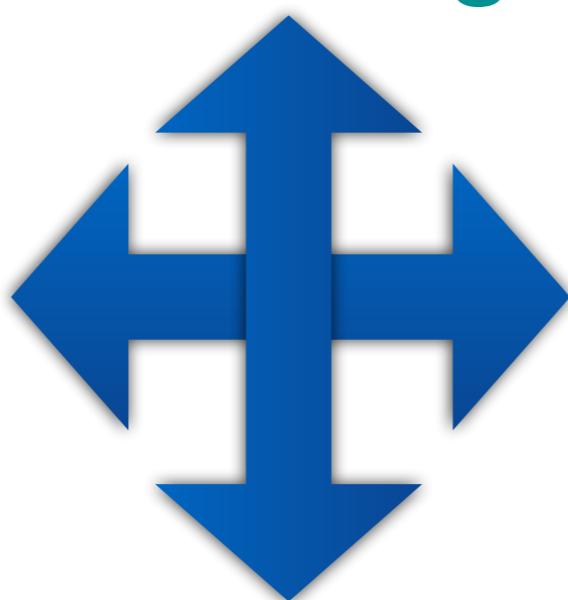
Application for diagnosis





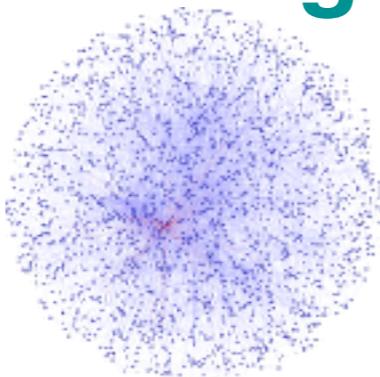
Dynamical Network Modeling

Algorithmic
developments



Applications

Large-scale Network
mining





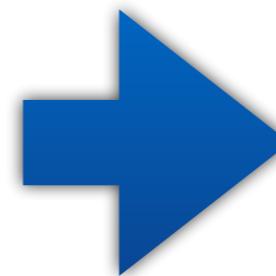
Cancer targeted therapy

- Combinatorial drug treatment / **Synergy**
 - Increase success
 - Decrease resistance
- 150 drugs => > 10.000 pairwise combinations
- Can we develop a **signaling pathway model** that can predict **synergies** ?

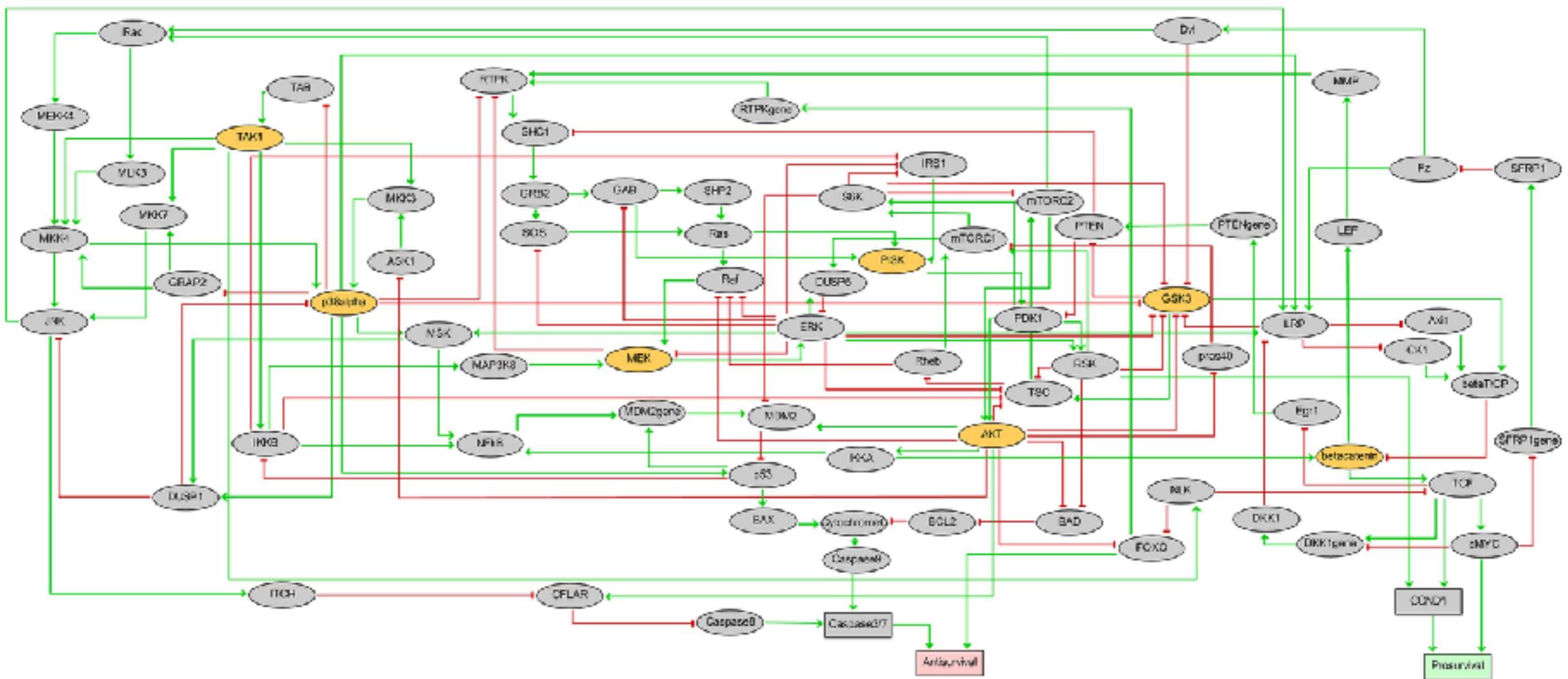


AGS gastric cancer cell line

- 4 signaling pathways
- 2 phenotypic outputs

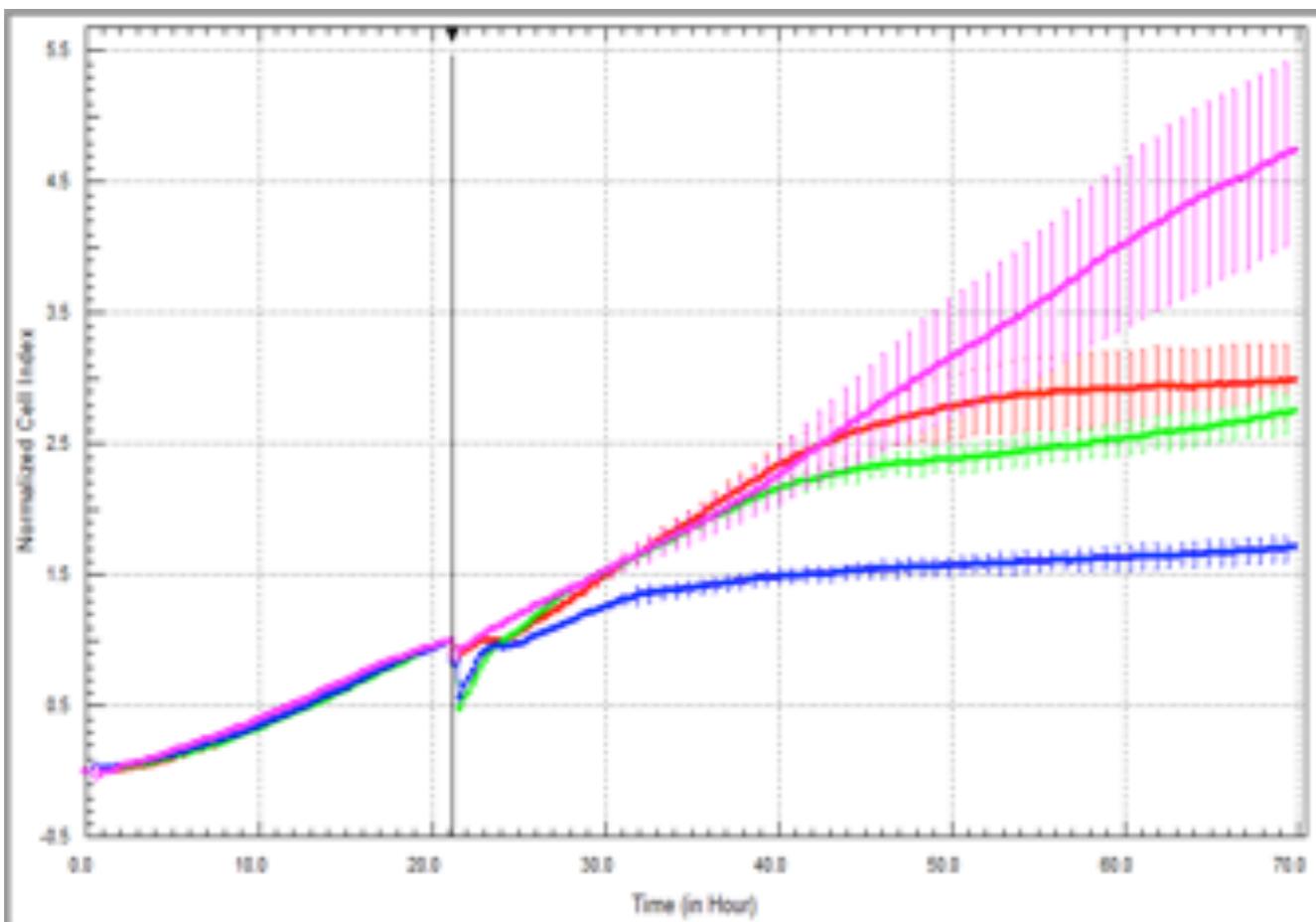


In silico screenings





Drug synergies



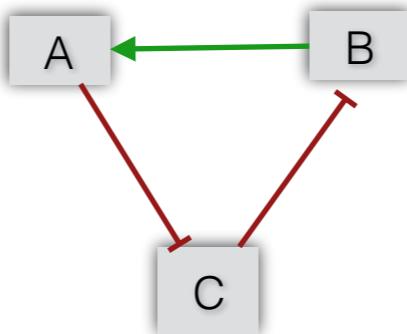
Control

Drug A

Drug B

Drug A + B

4/5 validated experimentally on cell growth experiment



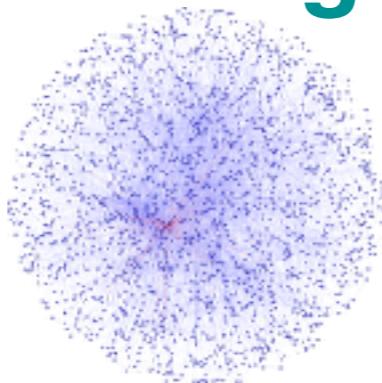
Convergence ?



Dynamical Network Modeling



Large-scale Network mining





PEPS Bio-Maths-Info



INSTITUT DE MATHÉMATIQUES de MARSEILLE

Elisabeth Remy
Laurent Tichit
Alain Guénoche
Gilles Didier
Brigitte Mossé
José-Luis Portero

Alberto Valdeolivas



Christine Brun
Andreas Zanzoni
Lionel Spinelli



Luc Camoin
Stéphane Audebert
Jean-Baptiste Boyer
Palma Rocchi
Maria Katsogiannou

UGMGF / ProGeLife

Pierre Cau
Claire Navarro
Sophie Perrin
Nicolas Lévy



Astrid Laegreid
Asmund Flobak



Alfonso Valencia
Kristina Ibañez
César Bullosa

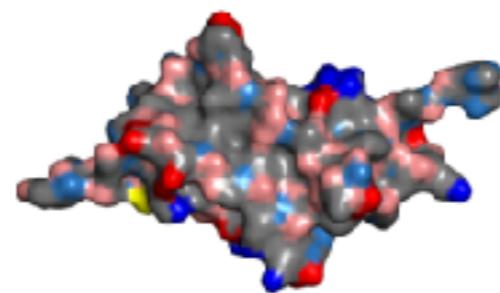


Rafael Tabarés-
Seisdedos



Hsp27 role in Prostate Cancer evolution to

Resistance



hormono-sensitive

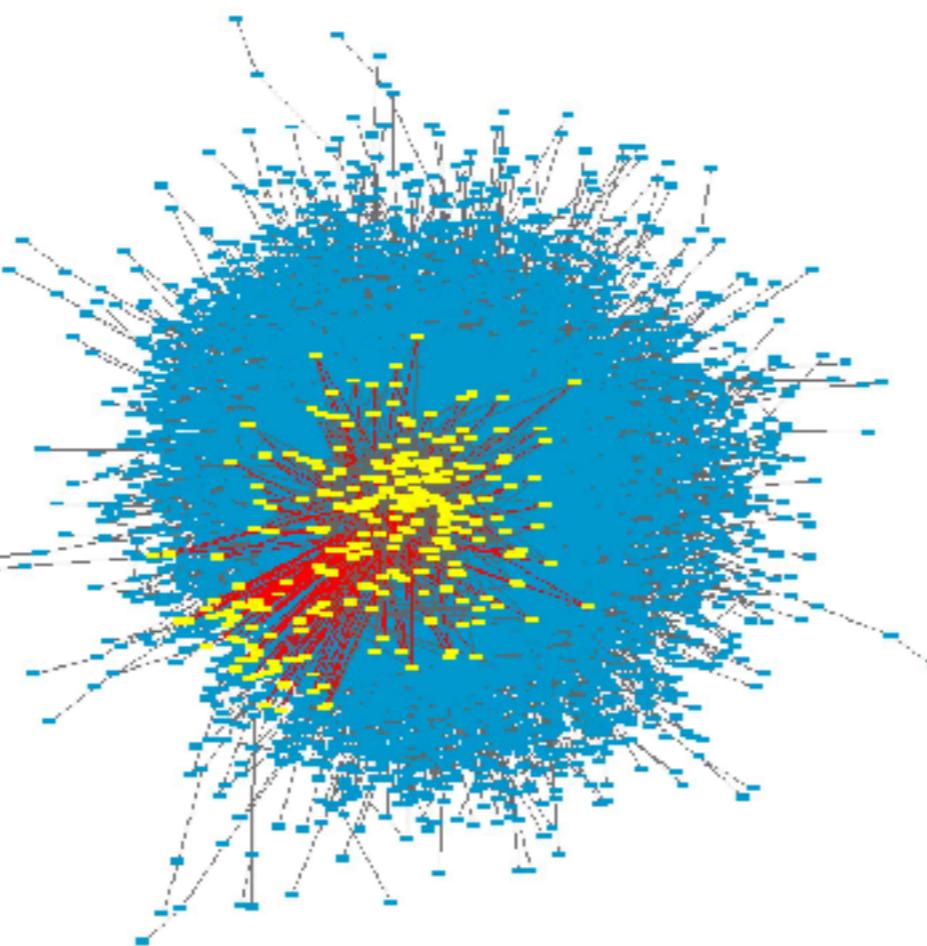
hormono-resistant

Hsp27 expression

Interactome mapping



226 interactors
for Hsp27

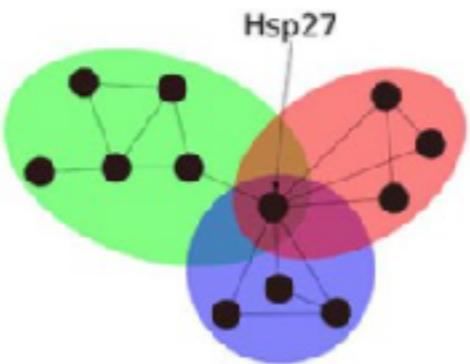


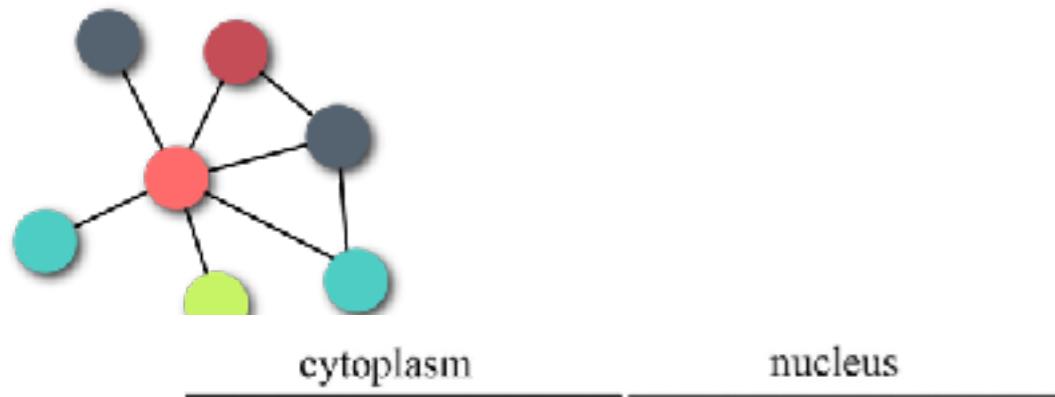
Human
interactome

80 877 PPIs
10 282 proteins

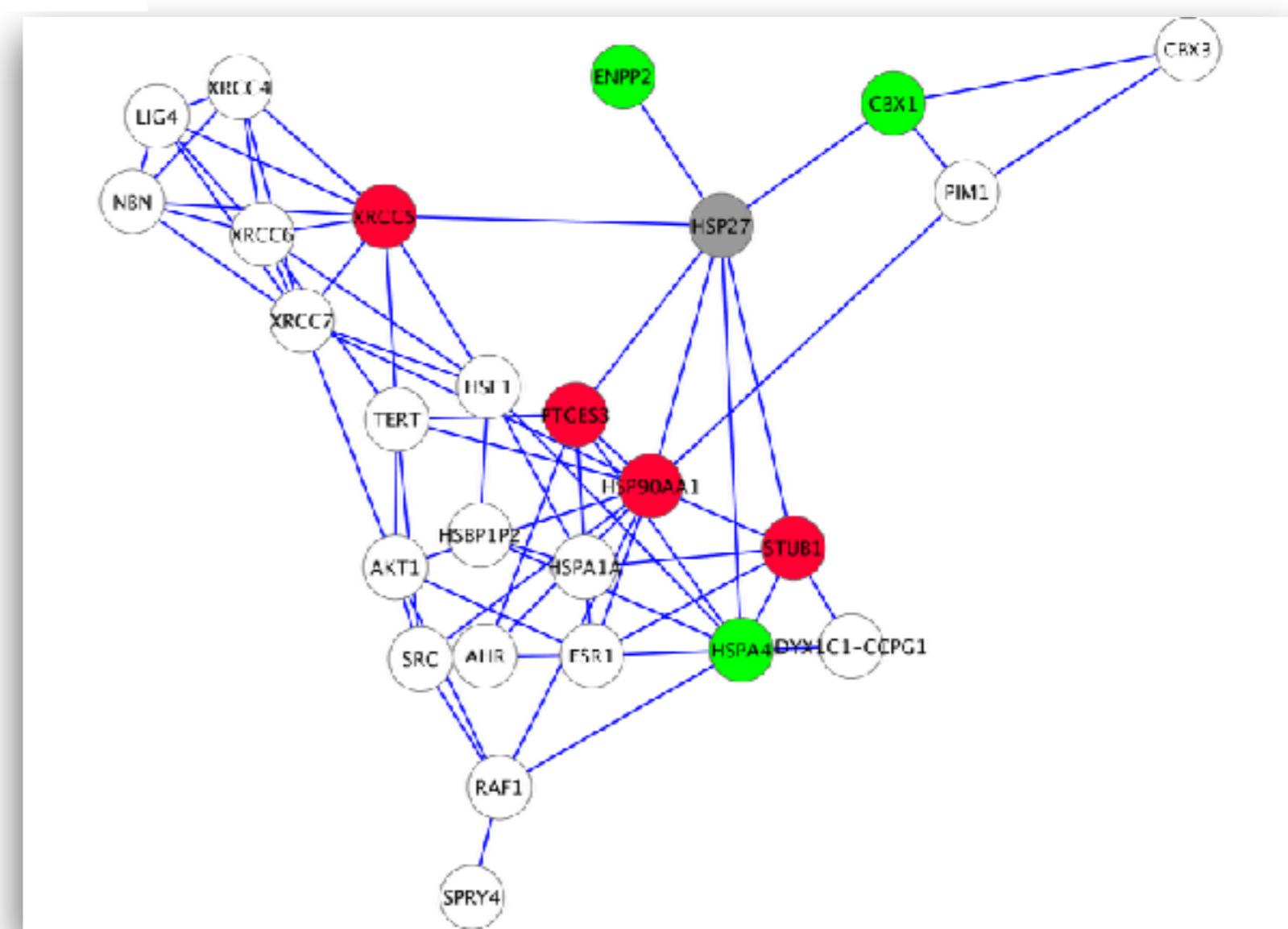
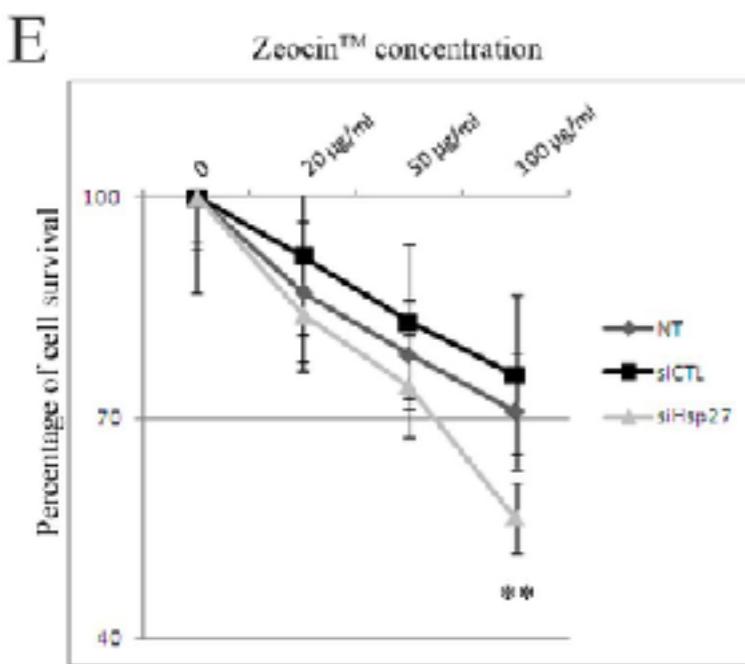
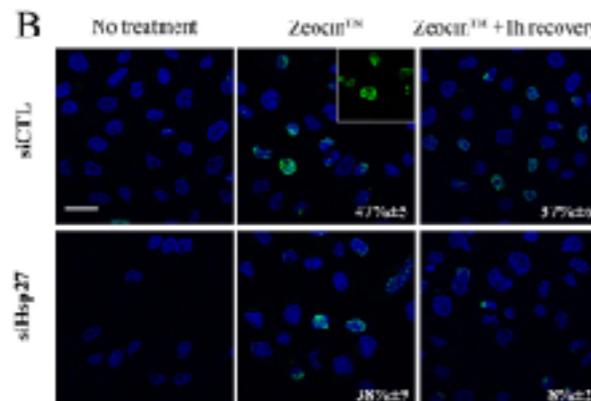
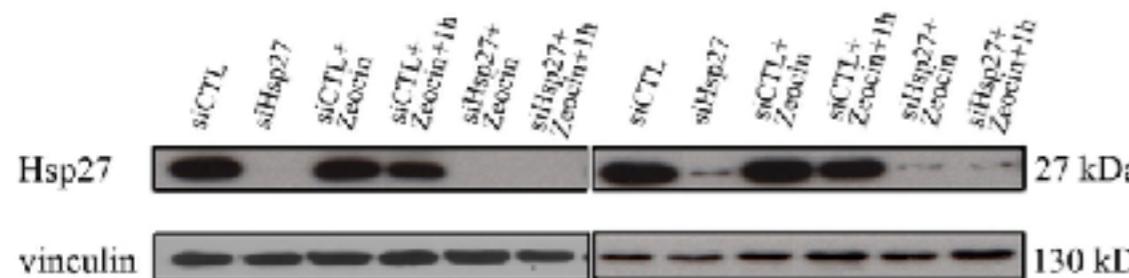
Module identification
(OCG algorithm)

54 Hsp27-containing modules
(over 751 overlapping modules)

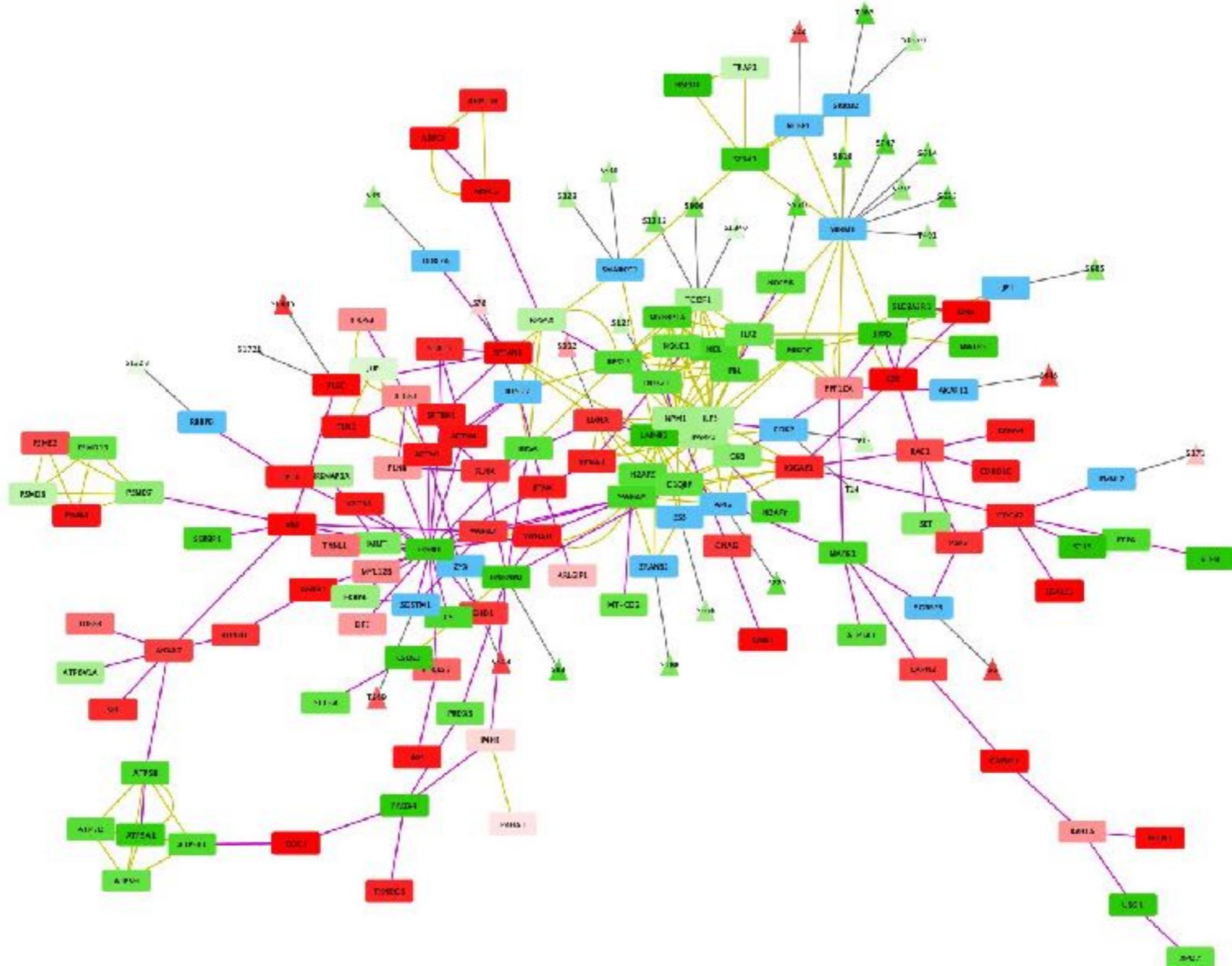


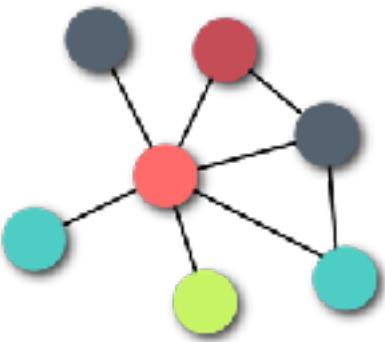


Hsp27 is involved in DNA repair



Proteomics/phospho-proteomics



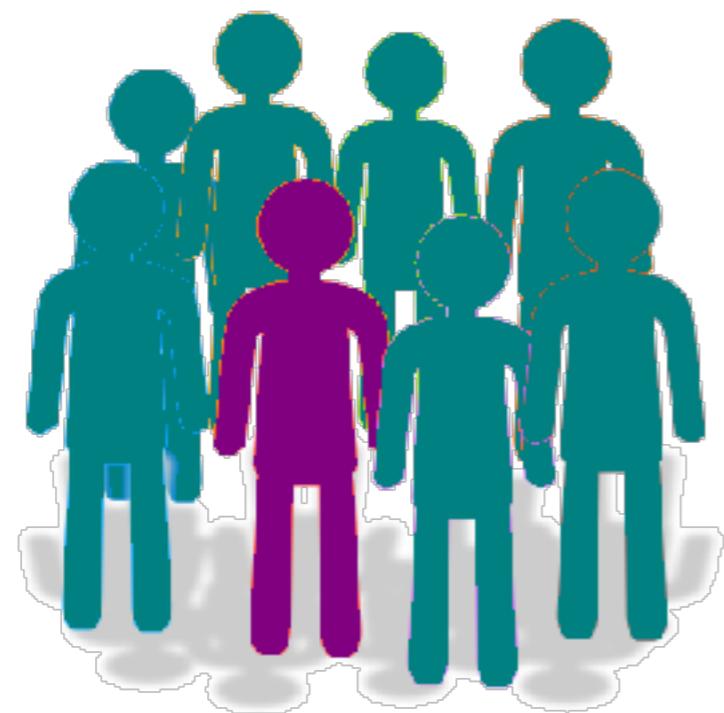


Disease comorbidity

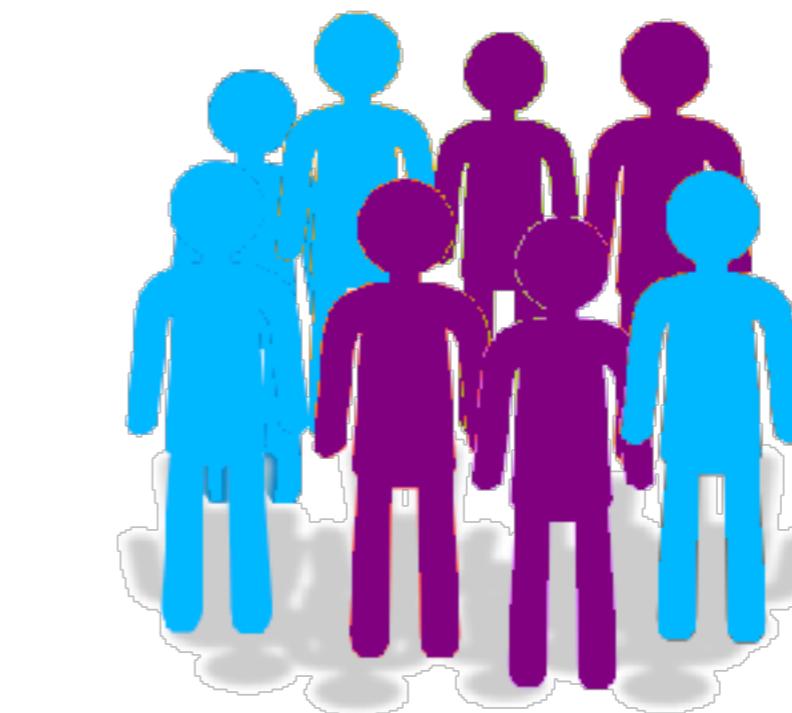
Comorbidity / Multimorbidity

Increased co-occurrence of diseases

Ex : Diabetes mellitus and cataracts



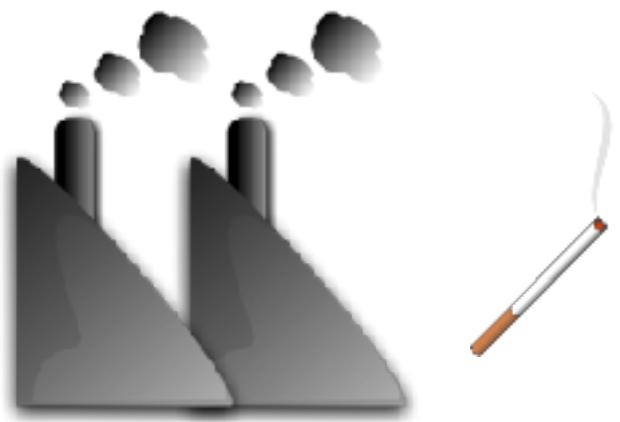
Control



Diabetes



Why?



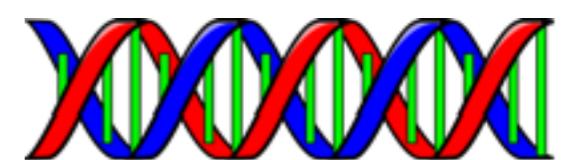
Environment
Lifestyle



Lower screening
survey



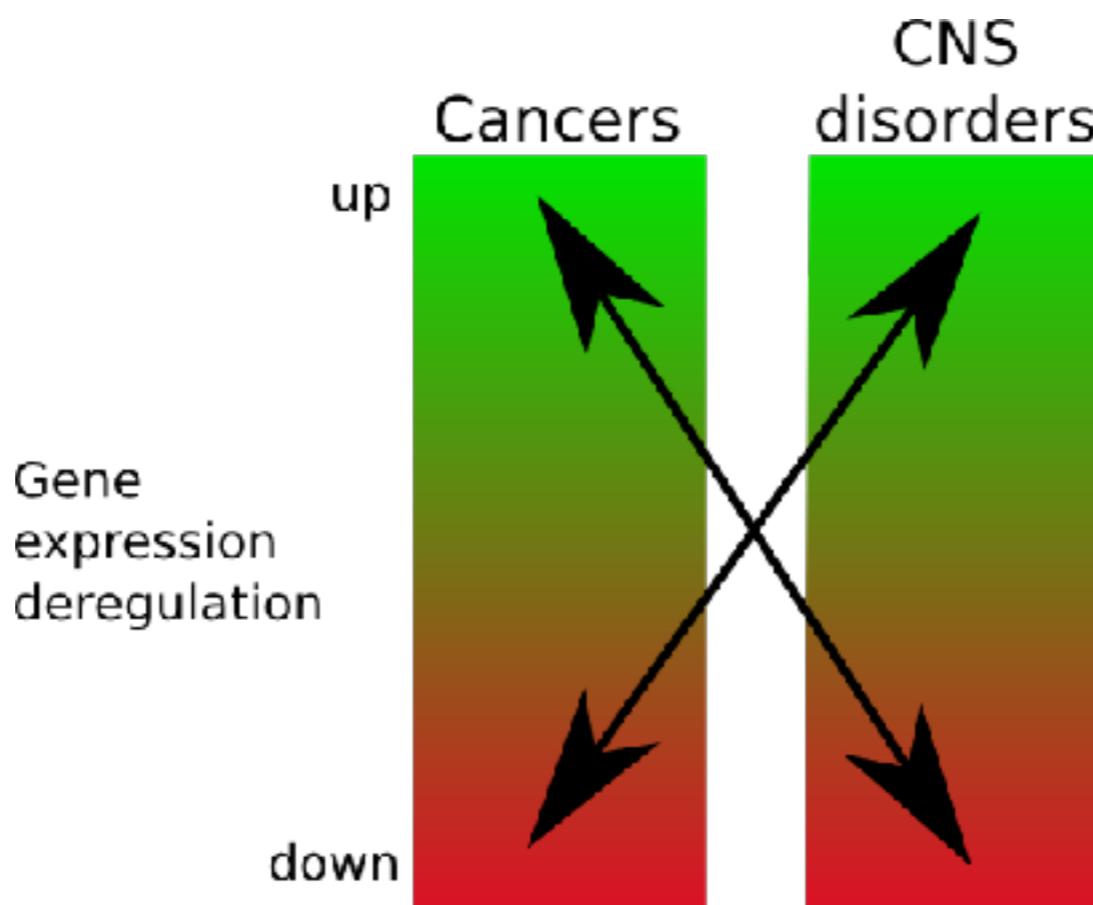
Drug treatments

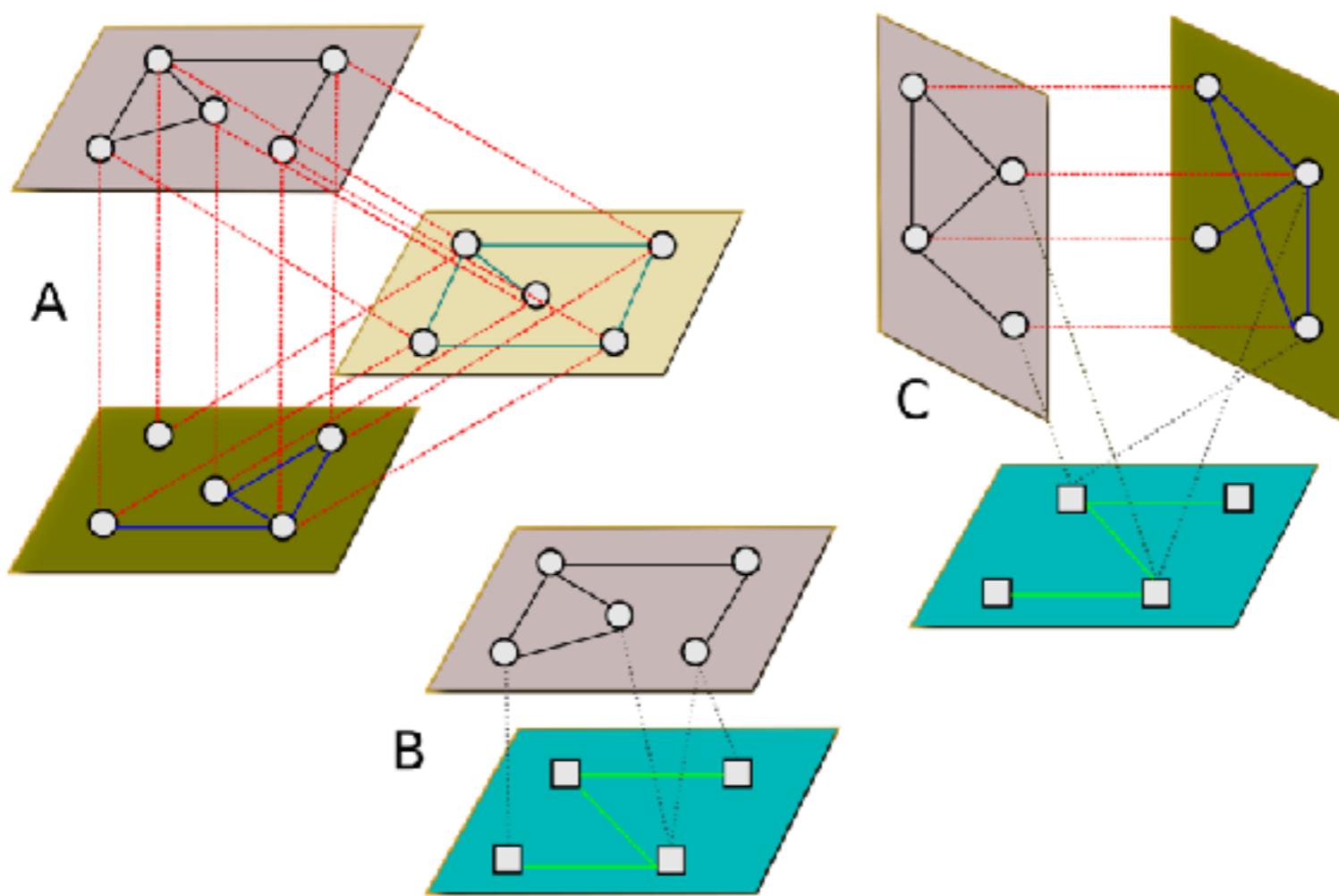


=> Genetics ??



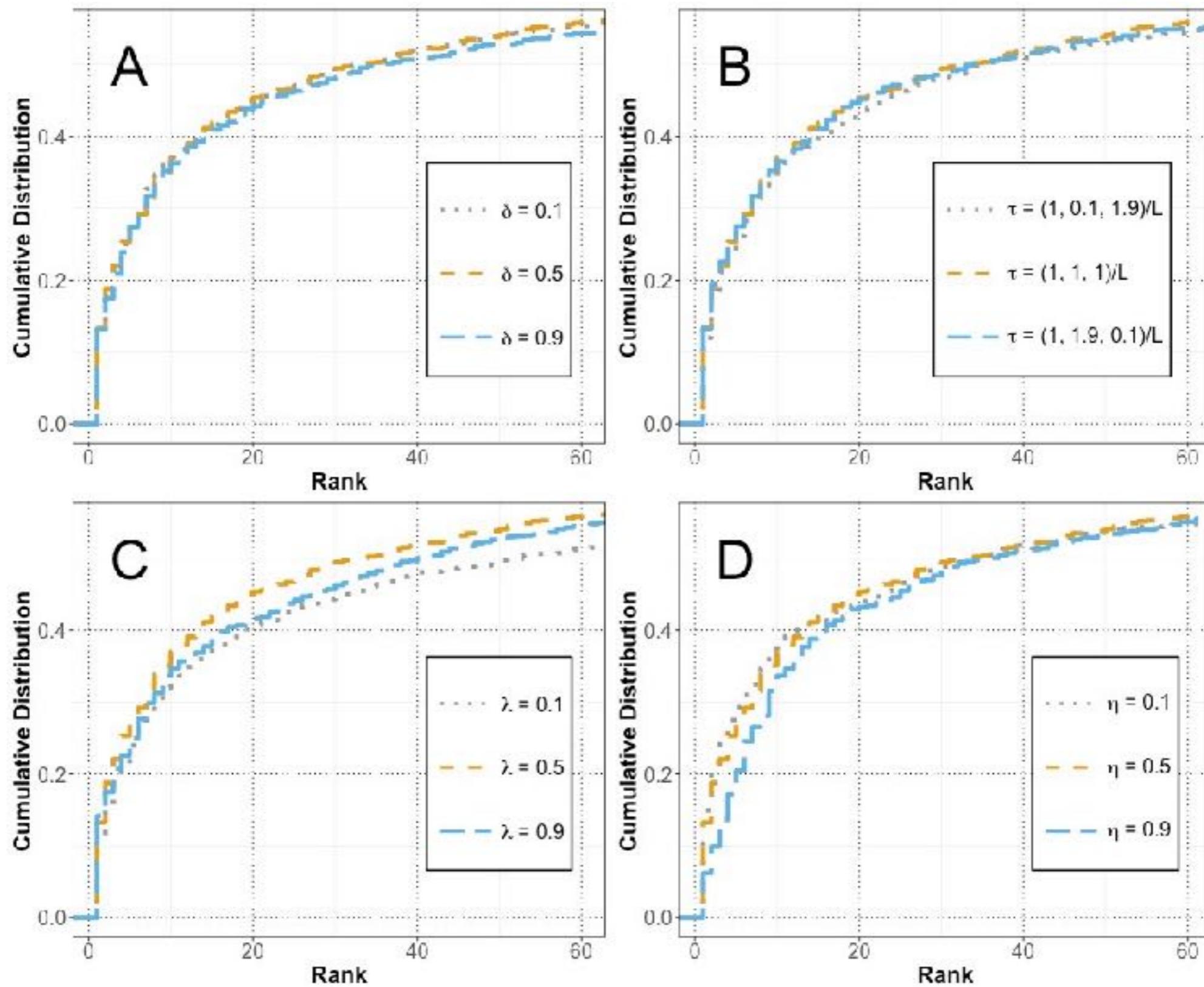
Inverse comorbidities / Transcriptomics meta-analyses



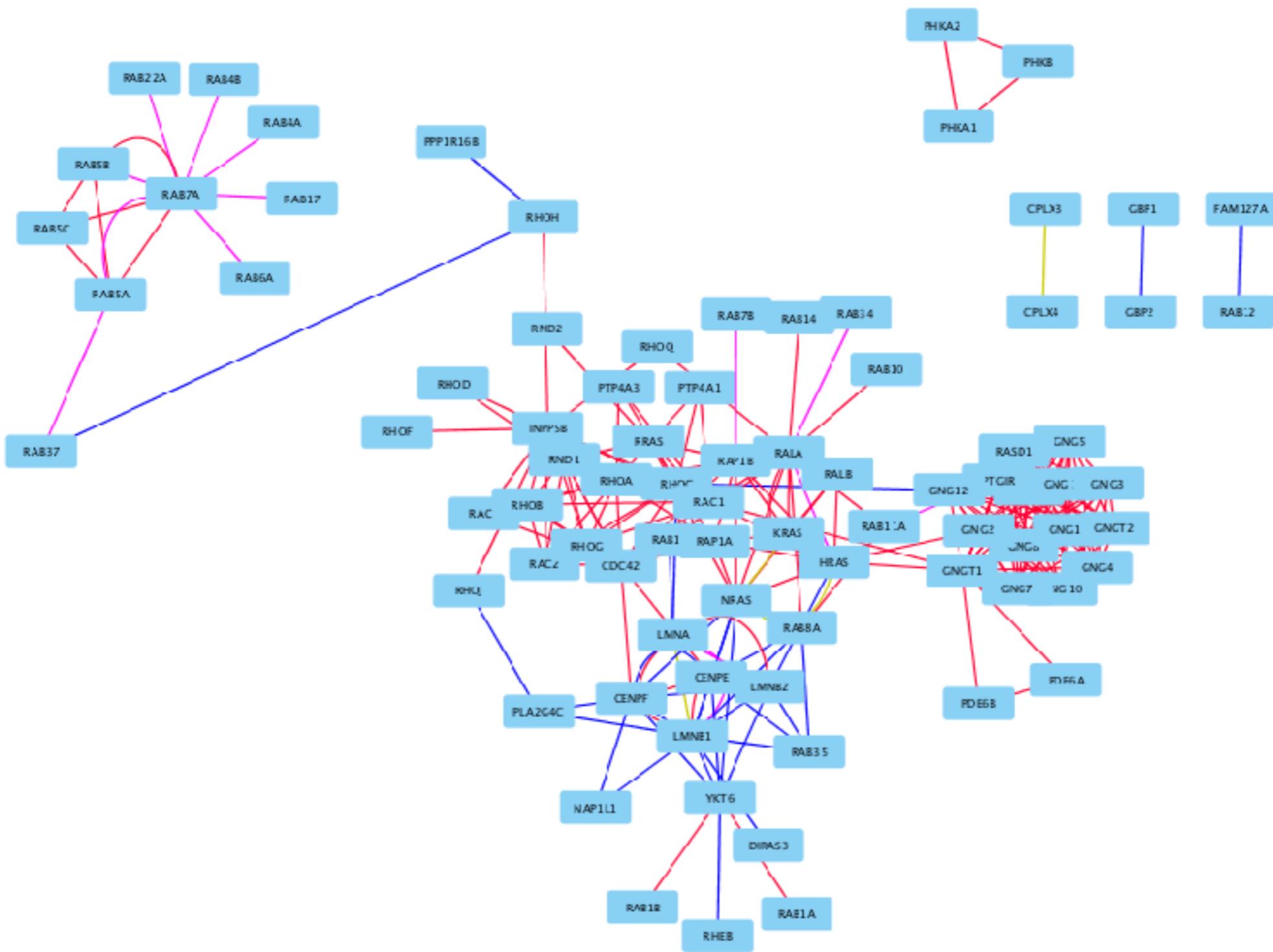




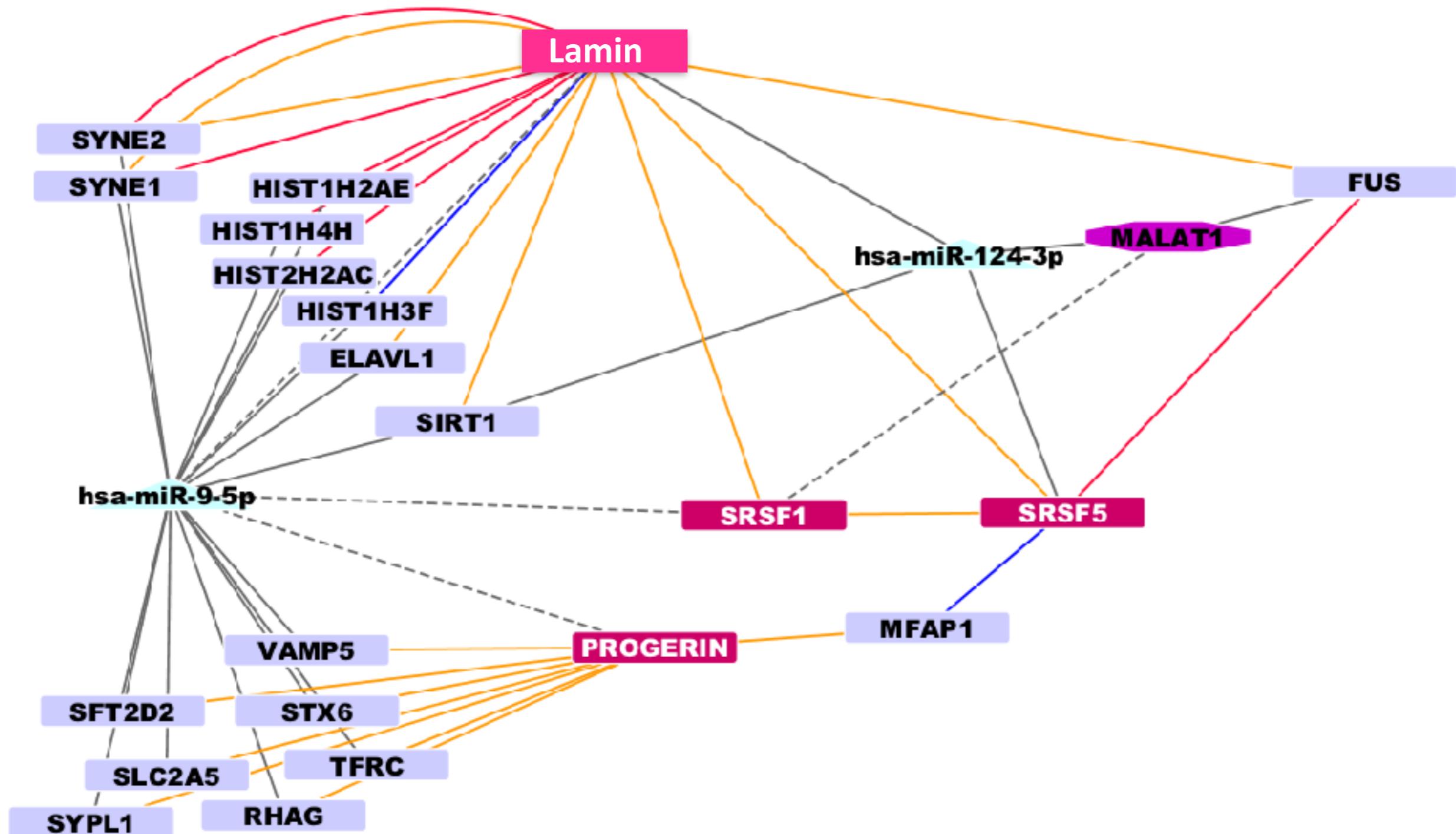
RWR-MH parameters



Prenylome preliminary analyses: Interactions between prenylated proteins



-omics data integration



Patrice Roll

Data integration, -omics algorithms

